5-1-2005

A comparative analysis between the RIT Contrast Resolution Test Target And the Gutenberg Test Target

Deepak Dubey

School of Print Media
Rochester Institute of Technology
Rochester, NY

Certification of Approval

Master's Thesis

This is to certify that the Master's Thesis of

Deepak Dubey

has been approved by the Thesis Committee as satisfactory for this thesis requirement for the Masters of Science degree at the convocation of

April 2005

Thesis Committee:

Dr. Edward Granger
Primary Thesis Advisor

Franz Sigg
Secondary Thesis Advisor

____
Thesis Consultant

Dr. Twyla Cummings
Graduate Thesis Coordinator

Dr. Twyla Cummings
Graduate Program Coordinator

Dr. Patricia Sorce
Chair, SPM

A Comparative Analysis between the RIT Contrast Resolution Test Target And the Gutenberg Test Target

By
Deepak Dubey

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in the school of Printing Management and Science of the Rochester Institute of Technology

May 2005

Thesis Advisor: Dr. Edward Granger
Research Advisor: Prof Franz Sigg

Permission to Reproduce

1. Title of thesis "**A Comparative Analysis between the RIT Contrast Resolution Test Target and the Gutenberg Test Target ".** I, Deepak Dubey, hereby grant permission to the Wallace Memorial Library of RIT to reproduce my thesis in whole or in part. Any reproduction will not be for commercial or profit.

   Date: May 21, 2005

2. Title of thesis "**A Comparative Analysis between the RIT Contrast Resolution Test Target and the Gutenberg Test Target ".** I, Deepak Dubey, prefer to be contacted each time a request for reproduction is made. I can be reached at following email address:

   Date: May 21, 2005

Acknowledgement

I would like to thank my advisors; Dr. Edward Granger, Prof. Franz Sigg, and Dr. Twyla Cummings for their guidance, help, and patience during the completion of this thesis.

In addition, I would like to thank all of the observers, who participated in the tests, as their results are a vital part of the research in this thesis.

Table of Contents

List of Tables

List of Figures

Abstract

In addition to color, the other two most important properties that determine quality of reproduction of a printing/imaging system are contrast and resolution. Contrast and resolution limitations of an output device partially define output quality. These limitations can be attributed to various factors such as the screening method used by the RIP, the image transfer method of the output device, the substrates used, and capabilities of the postscript interpreter or a combination of all these factors.

This study introduces a new test target for quality evaluation called the Gutenberg Test Target. A comparative analysis between the Gutenberg Test Target and the RIT Contrast Resolution Target has been performed. Both targets have been developed collectively by Dr. Edward Granger and Franz Sigg and both are used to measure the contrast and resolution limitations of an imaging system.

The Gutenberg target uses a visual subjective comparison to evaluate overall output quality of an imaging/printing system. On the other hand, the RIT Contrast Resolution Target uses a method of analysis to calculate a Contrast-Resolution-Volume (CRV) value, which then is used as a relative indicator of the reproduction quality of the imaging/printing system.

For this study both test targets were printed on a selected imaging /printing system with varying degrees of image deteriorations (Gaussian Blur) applied on the images. The image deteriorations simulated different Modulation Transfer Functions (MTF) for different devices. The printed test targets were analyzed by performing two visual experiments, one for each target, using a number of observers. The generated data from the experiment was used for mathematical analysis and a comparison was made between the two targets.

The final analysis and results showed that both targets do a good job in measuring the resolution contrast limitation of the system.

Chapter 1

Introduction

In the past, extensive research has been conducted in the graphic arts industry to identify various quality parameters that can determine capabilities of printing/imaging devices by means of test targets, but relatively limited work has been done on determining performance of a printing/imaging system as a whole.

A printing/imaging system is a set of various stand-alone devices connected by a single workflow; examples of stand-alone printing/imaging devices are film-setters, plate-setters, proofing devices, or (digital) presses. Depending on the configuration of the devices used for the system, a printing/imaging system could be divided into three categories: digital, analog, and hybrid. A digital system is one which only uses digital devices in the workflow; an analog system uses only analog devices such as analog film preparation, plate making and offset lithographic press, whereas a system using both kind of devices would be a hybrid printing system.

The tools which are used to determine the output quality of a given printing/imaging device or of a system are known as Test Targets. A Test Target can be used to determine the capability of individual devices as well as the comprehensive capability of a system as a whole. When processed through a device or a system, a test target provides useful information about the various parameters of the device/system. The information generated is further used for optimization, calibration, characterization, and process control to extract the best possible output from of the process.

There are two types of test targets, natural images or synthetic, patterns of lines, circles, halftone patches which today are digitally programmed (PostScript).

Each element in a target is capable of testing one or more characteristic of the device or system. While some of the elements/ targets are subjective, i.e. can be evaluated visually or by any other subjective assessment method, other targets require the use of measuring devices such as a densitometer, colorimeter or a spectrophotometer. Such targets are called objective targets. Some examples of test targets are parallel lines, checkerboards, and star target (to measure directional effects), and registration target, natural SCID test images as defined by ISO 12640, solid area patches (ink density), tints and overprints.

Test targets can be used at any stage of a printing system. They are used to measure and control various output quality parameters such as color, color gamut, density, tonal quality (tone reproduction and details in highlight, shadows and middle-tones), line quality (width, blurriness, and raggedness), text quality and effective resolution.

Most of the digital test targets available in the graphic arts industry are used with output devices such as film-setters, plate-setters, proofing devices, digital presses, and desktop printers.

Image Quality reproduced by imaging systems is a direct function of their addressability. Output devices are generally divided into two categories depending upon their addressability: low and high addressability imaging/printing systems. The resolution of a system is defined as its ability to reproduce fine detail of an image[1] that is a function of the output device's mechanical ability to image small spots (addressability) and its ability to render density difference between the foreground and the background (contrast).

Different devices use different methods to generate detail and contrast to reproduce an acceptable quality of image. A high-addressability imaging device uses its high addressability, i.e. the mechanical ability of the system to put very small spots inside a halftone cell, to generate grey levels.

On the other hand, a low-addressability imaging device has a smaller number of spots within a half tone cell and therefore, does not have the physical ability to generate enough gray levels to produce a good quality image. To compensate for this shortcoming low addressability imaging devices can image a given spot at various color levels, utilizing bit depth or multilevel inking ability, and therefore, can produce image quality comparable to high addressability devices. The number of gray levels at which a single spot can be imaged is expressed as a binary count and is referred to as bit depth.

A system under any given circumstances reproduces the finest detail at maximum contrast (100%, only fully black or clear image areas) because no halftones are needed. As the contrast between the detail and the background becomes less than 100%, a screening method becomes necessary to simulate the density difference. The simulation generated by the screener/screening method causes loss of resolution because of the halftone dots. As the density difference is further reduced, after a certain point, the system becomes incapable of maintaining any kind of density difference/contrast between the detail and the background. Such limitations are caused by interactions between various frequencies, which are the screening method, the addressability of the output device, and the resolution (ppi) of the original image.

Knowing the resolution contrast capability of systems available can help in making a better decision about whether the system is suitable for a particular job. It can also help in using the system to its maximum capability to extract optimum results.

This thesis employs two test targets, the RIT Contrast Resolution Target and the Gutenberg Test Target, developed by Dr. Edward Granger and Prof. Franz Sigg, for measuring the contrast and addressability limitations of imaging systems/devices.[2]

The RIT Contrast Resolution Target, along with its method of analysis, calculates a contrast-resolution-volume (CRV) value for the system. Elliot Harper, in his master's thesis, proved that the RIT Contrast Resolution Target provides a method of discriminating CRV of marking engines and screening methods.[3]

This thesis introduces a new test target called the Gutenberg Test Target. The Gutenberg Test Target has been collectively developed by Dr. Edward Granger and Prof. Franz Sigg for measuring the resolution (print quality) limitation of an imaging system/device.[4]

The purpose of this thesis is to perform a comparative analysis between the RIT Contrast Resolution Target and the Gutenberg Test Target, to determine whether the Gutenberg Test Target can discriminate between different printing systems and whether it provides an easier method for evaluating print quality than the RIT Contrast Resolution Target.

Endnotes for Chapter 1

1. Franz Sigg, "*Definition of Imaging Terms (5th draft)*" (Rochester, NY: RIT, 2003).
2. Franz Sigg, "*RIT Contrast Resolution Test Target*" (Rochester, NY: RIT, 2000)
3. Eliot Harper, "*An Investigation into the Relationship between Contrast and Resolution*" (Masters Thesis, Rochester, NY: RIT, 2000).
4. Franz Sigg, "*Workflow to make Gutenberg target*" (Private communication, Rochester, NY: RIT, 2002).

Chapter 2

Theoretical Bases of Study

Introduction

In order to understand and interpret the contrast resolution differences between different printing systems, it is first necessary to identify the underlying principles and mechanics behind imaging/printing processes. This chapter provides an explanation and identifies the fundamental differences between contrast and resolution and the effect of the screening method, and explains the structure and evaluation methods of the RIT Contrast Resolution Test Target and the Gutenberg Test Target.

Before any further discussion, it is important to understand some basic concepts, definitions, and terminologies regarding contrast and resolution.

Definitions

Bit Depth

Bit Depth is defined as the number of bits of tonal range capability of the spots of an output device. For example, a graphic arts film or plate setter has a bit depth of one, i.e. each spot can only be either on or off. On the other hand, bit depth for a display device is the number of bits of tonal range capability of the pixel in an image. For example, RGB 24 bit color means a pixel depth of 8 bits per color or 2^8 (256) levels per color[1].

Dots

In terms of printing, a dot is the smallest screener element. For a binary output device having high addressability and a bit depth of one using amplitude modulation (AM) screening, a dot is composed of many spots. This dot is commonly known as a halftone dot and its fineness is measured in "lines per inch" or lpi. In the case of a display device that has a large bit depth (8 bits or 256 tone values) at a low addressability (72 spi.), a dot has the same size as one spot.

Marks/Spots

Marks are the physical elements that the output device or the marking engine places at the addressability locations that in turn are called spots. A mark can have a different size and shape than one spot (addressability location). The size of the spot is determined by the addressability of the output device. A typical computer monitor has 72 spots per linear inch while a film imagesetter may have more than 4000. A typical digital press prints at 600 spots per inch with a few models printing at 800 or more per linear inch. The "spot" is often confused with the "dot" as in "dots per inch", but a printed dot is actually made up of a group of spots. It is common to refer to the addressability of an output device as dots per inch (dpi), even though it actually means spots per inch.

Marking Engines

Marking engines are those physical devices that place marks on the addressability grid on the output medium[2]. The term is usually applied to imagesetters, platesetters, and digital printers.

Raster Image Processor (RIP)

An RIP is a computer program that takes the input information from an image and generates a spot map, i.e. patterns of spots arranged as a bitmap which creates a visual representation of the original image on the output device[3].

Methods of Transferring Images

There are two main methods used in transferring images to the substrates:

1. Binary system.

2. Bit Depth system/ Multilevel inking.

Binary Method. Output devices such as imagesetters, platesetters and many digital printers / proofing devices are called binary devices because they use only one or zero to control the on/off state for each spot. In this method, a halftone cell is made up of many spots and the different levels of gray are achieved by turning some pixels on or off within a halftone cell. Systems employing this method need high addressability in order to achieve enough gray levels to represent continuous tone images at high resolution.

Bit Depth / Multilevel Inking Method. These systems are able to generate different tone values by applying varying amounts of ink on a given spot. Systems having low addressability have multilevel inking capability to produce the levels of grey required to reproduce high quality images.

Screener

It is that part of RIP which calculates area modulation or multilevel inking values of each spot for an output device, on the basis of required tonal value of the projected pixel from the input image, in order to form halftone dots. The function of a screener is very device-dependent. The screener tries to render tone values that represent the information of the input pixel grid[4].

Screening Method

These are the methods used by the screener to form halftone dots on the addressability grid of the output devices with a bit depth of 1 (binary). There are two types:

1. Amplitude Modulated Screening (AM).

2. Frequency Modulated Screening (FM).

AM Screening

FM Screening

Figure 1. *Examples of AM and FM screening*

Amplitude Modulation Screening (AM). In this screening method, different dot sizes are achieved by turning spots on or off within a halftone cell. The distance between the centers of the halftone dots is constant; the placement of spots within the dot is clustered; that is, spots are centered within the halftone cell to produce different tonal values.

Frequency Modulation Screening (FM). The dot shape and size are kept fixed but are placed randomly on the addressability grid of the output device. The placement of spots within the halftone cell is dispersed; that is, spots are placed randomly within the halftone cell to produce different tonal values.

Gaussian Blur

A Gaussian probability curve is the classic bell-shaped curve with a higher probability in the center of the curve and lower probabilities on either side. A Gaussian blur effect takes each pixel and mixes it with adjacent pixels with Gaussian probability so that the pixel has greater effect near its original location and lesser effect (in a bell curved shape) farther away from its original location[5].

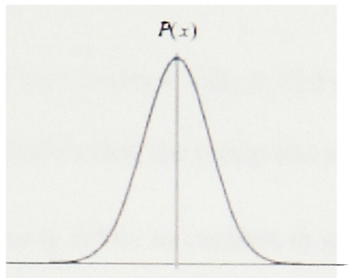


Figure 2. *Bell Shaped Probability curve*

Contrast

The term refers to the visual difference between two adjacent parts of an image. Black and white colors have the greatest contrast. In an image, the density difference between the lightest and the darkest area characterizes its tonal range with the lightest and the darkest points as the endpoints. An image is considered to be high contrast if there is a large density difference between its lightest and its darkest regions. Figure 3 is an example of a high contrast image.

Figure 3. *High contrast image*

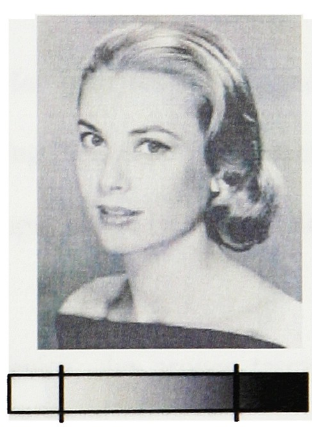


Figure 4. *Low contrast image*

The scale below the picture shows the range of gray scale used between the lightest and the darkest image areas; the large gap between the two points on the scale shows that the image has a high contrast.

Figure 4 is an example of a low contrast image. The smaller gap between the lightest and the darkest image areas indicates that the image has a low contrast.

The above illustrated example only refers to contrast in grayscale images; differing levels of contrast can also occur in colored objects.

Resolution

Resolution is a term that describes the capability of an imaging system to reproduce small detail of an image.

People often confuse resolution and addressability. Unlike resolution, addressability is strictly a fixed, mechanical characteristic of the output device. Resolution on the other hand relates to the subjective perception of an image. Resolution cannot be a single number because it is a function of contrast. More resolution is obtained at high contrast than at low contrast.

Addressability is one of the contributing factors for both resolution, and gray levels (which in turn relates to contrast). Other contributing factors are the screening method, registration of the colors and the sharpness of the input image (which is related to image processing and the number of pixels of the image).

When describing spatial aspects of an output device, the proper terms to use are addressability and possibly screen ruling, but not resolution.

There are different quantitative units affecting resolution; pixels per inch (ppi), samples per image (spi), and dots per inch (dpi). The following definitions of these units are used in this report:

Pixels per Inch (ppi)

A pixel is the smallest picture element of a digital image. Pixels are used to indicate the total amount of information that a digitized image contains in its horizontal and vertical dimensions. The term "pixels per inch" refers to the number of pixels in an inch contained in a digital image. It is used to measure the total amount of raster information within an image. For a display device, resolution defines the number of discreet horizontal and vertical visual elements. It is also called screen resolution. Pixels are data, not a physical element, and refer only to the input side of an image. On the output side, there are no pixels, but spots or halftone dots.

Samples per Image (spi)

A sample per image is the correct term for referring to the input resolution of scanners and digital cameras. For such devices, the term pixel per inch refers to the sampling rate of the device and the amount of pixel information in the image. Input devices offer a variable range of resolutions. As the sampling rate of an image increases, the pixel size decreases, thus enhancing the image resolution. This is illustrated in Figure 5.

50 ppi 100 ppi 300 ppi

Figure 5. *As input resolution increases, pixel size decreases*

Dots per Inch (dpi)

This term is used somewhat loosely by the industry. Normally, people use this term to describe the addressability of an output device. However, an output device images spots not dots. The term dot properly refers to halftone dots. In AM screening, a halftone dot is placed at the center of a halftone cell, and consists of a number of spots. The correct unit for addressability is spots per inch (spi) and not dots per inch (dpi).

Halftone Cell (dot matrix)

A halftone cell is made up of a grouping of spots to form a dot structure. The size of a halftone cell is defined by the number of spots along its x and y-axis and determines the tonal range it can achieve. A halftone cell uses all the possible spot positions to construct a dot. Figure 6 shows a 2X2 halftone cell. In this cell, the possible conditions for filling the halftone cell are no spot, one spot, two spots, three spots, and four spots.

In a halftone cell, the discreet tonal values (or gray levels) are restricted by the maximum number of spot positions available within the cell, plus one (no spot, blank). Hence, for a cell size of 4X4, the maximum number of spots would be 17 (4X4+1).

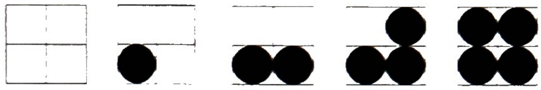


Figure 6. *The possible combinations for filling a 2X2 dot matrix*

A halftone dot uses all the possible combinations of spots to achieve different tone values. These types of marking combinations are known as perturbations. Figure 7 shows the various possible combinations of spots within a 2X2 matrix.

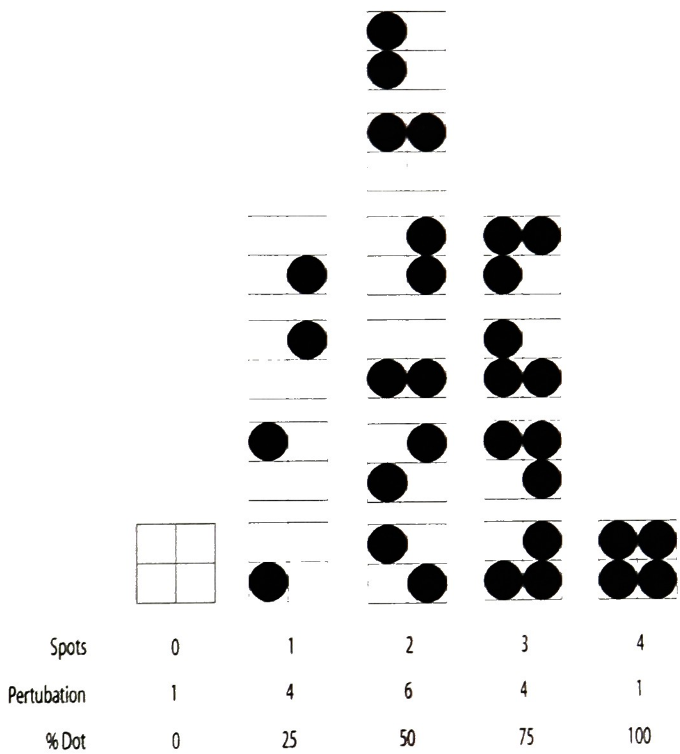


Figure 7. *The 16 possible perturbations for a 2X2 dot matrix*

As shown in figure 7, for each number of spots, there are several perturbations available. The exceptions are zero spots and maximum spots, where only one possible perturbation is possible. Although such groupings are equal in dot area coverage or percentage density, they vary in dot structure. Therefore, a 2X2 cell has a minimum of 5 gray levels and has a total number of 16 spot configurations.

Screening Methods

One of the main factors which control the contrast and resolution capabilities of an output device is the type of screening method used. Generally, there are two types of screening methods used in graphic arts today, AM screening and FM screening. This section explains each screening method and notes its effect on contrast and resolution.

Ordered Dither

Ordered dither is a method of providing a fixed sequence of turning pixels on within a halftone cell to generate the gray levels. There are two main approaches; clustered spots and dispersed spots.

Clustered spots. At a given tone-level, clustered-spot-ordered dither turns adjacent pixels' spots on, forming a cluster in the cell. This grouping of pixels results in substantial low-frequency components. The dot-to-dot spacing is fixed and the tone level is modulated by dot size within the halftone cell. This technique is also called the "amplitude modulation" (AM) technique.

Amplitude Modulated (AM) Screening

In four color AM screening, the RIP converts the image into four separate C, M, Y, and K images, and applies an amplitude modulated halftone screen pattern to each channel. The halftone dots in each screened channel are arranged at a given angle. The amplitude modulated screening is described by a set of parameters consisting of a screen angle, screen frequency, dot pattern, and level assignment (tone).

Screen Angle. Using AM screening for printing more than one color, the separated color screens must be rotated to prevent screening interference patterns, such as moiré or rosettes. The ideal angle between two colors which reduces interference patterns to a minimum is 30°. For color printing, the conventional approach for screen angles is to set black at 45 °, cyan and magenta at ± 30 °, and the yellow at 0 ° (Figure 8).

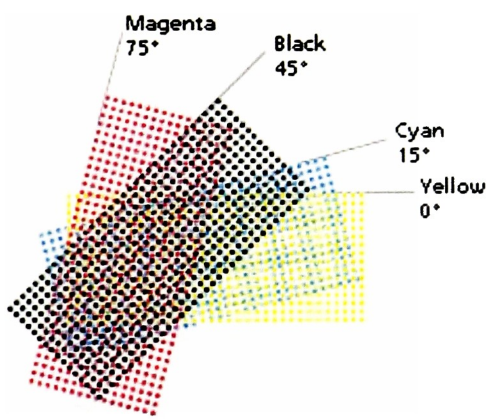


Figure 8. *Screen angles used for conventional four color printing*

Screen Frequency. It is a measurement of the number of halftone dots within a given unit of measurement. This is commonly given as a measurement of lines per inch (lpi) or lines per cm (cm^{-1}). By increasing the screen frequency, the dot size decreases, and the halftone frequency pattern becomes less apparent within the image. A low frequency gives a coarse appearance whereas a high frequency produces a fine and smooth appearance. Also, by increasing the screen frequency the number of tone levels achievable by a halftone dot gets reduced.

Two important factors in considering screen frequency are the image content and device capability. If the image contains areas of fine detail, then a coarse screen would be inappropriate because detail would be lost. However, a high screen ruling, which means small halftone dots, reduces the number of gray levels that can be imaged. Therefore, a compromise must be found between resolution capability and gray level capability

Dot Pattern. The dot pattern is the fill-in order for the dot growth sequence and is directly related to image texture. It has a strong impact on the output quality of an image. A digital halftone dot cannot grow uniformly as in conventional screening because of the fundamental discrete nature of the addressability grid. The dot can only grow by adding one printer spot at a time. In doing so, it is easy for the dot to become lopsided or asymmetrical, producing visual artifacts that show up as undesirable texture and coarse patterns in the printing.

Level Assignment. The level assignment relates to dot area or tone value.

Within the family of clustered-dot-ordered dither methods, there are four main techniques and variations for AM halftone dot generation. These are rational tangent, super cell, and multi-center dot.

A digital half-tone cell is constructed in a pixel grid. Depending on the relative positioning of the cell with respect to the grid, angled screens are classified as rational and irrational tangents.

Rational Tangent

A screen is called a rational tangent if the four corners of the halftone cell fall exactly on intersection points of the pixel grid[8]. It is shown in Figure 9.

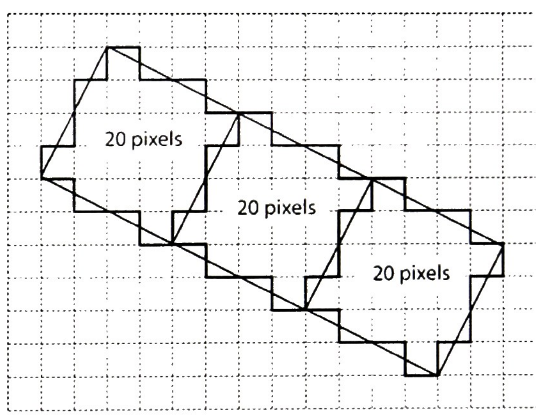


Figure 9. *Rational Tangent dot placement on a digital grid*

Rational tangent cells have exactly the same size and shape; they line up the same way with respect to the digital grid. Only one screen function is needed for all halftone cells. However, the match of cell corners to the digital grid limits the number of choices available for angles and frequencies. This makes it difficult to avoid moiré and artifacts unless large halftone cells are used by printing at coarser line screens.

Super cells

Angles of particular interest such as 15° and 75° cannot be accurately produced by rational tangent screens if the cell size is small. The problem is reduced by using large halftone cells, having large integers for the tangent ratio, which allows for smaller angular increments. It is therefore possible to achieve close approximations at 15° and 75° angles. The solution to this problem is to divide the large cell into many smaller sub cells to increase the screen frequency[9]. Using this technique, accurate angles can be produced. This approach is called super cell screening and is shown in Figure 10.

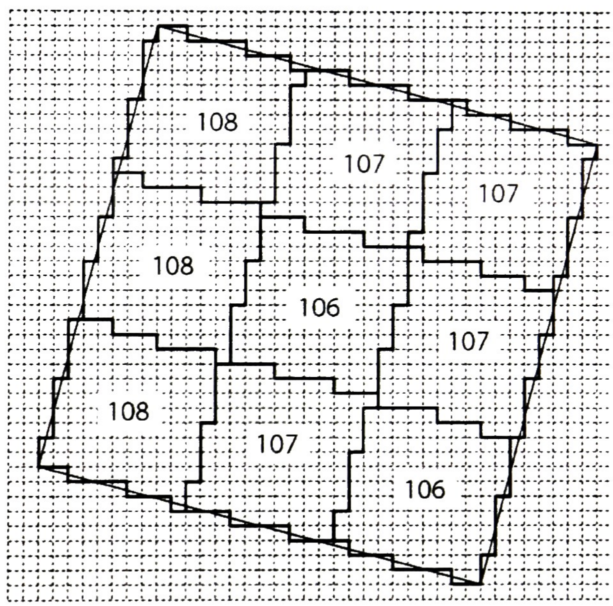


Figure 10. *Super Cell Irrational Tangent dot placement on a digital grid*

A super cell is a rational tangent screen composed of many smaller sub cells which are only approximately uniform in shape and size.

Multicenter dot

As shown in Figure 12, a multicenter dot cell is divided into approximately equal parts usually two or four. Each partial dot has a nucleus and grows into a separate cluster. The purpose is to increase the apparent screen frequency without reducing the number of tone levels[10]. An example is shown in Figure 11 where a 40-level dot is divided into four sub cells with 10 spots each.

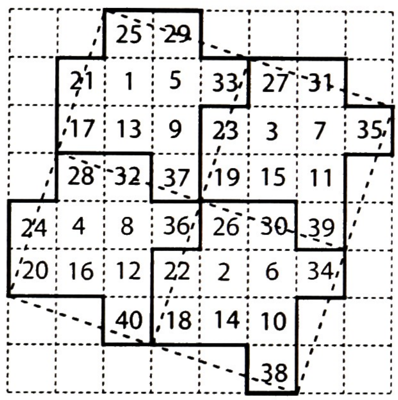
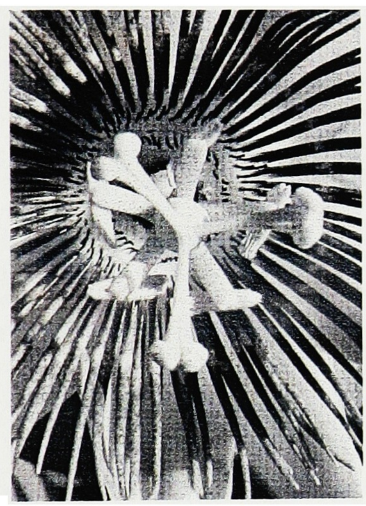


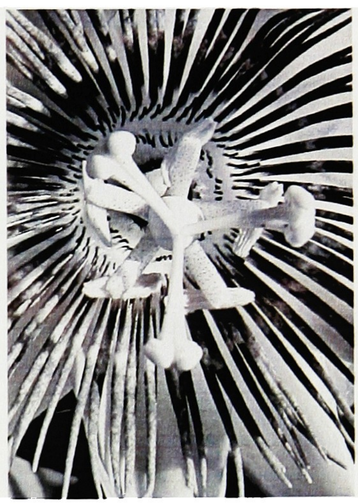
Figure 11. *Multicenter Dot Super Cell Irrational Tangent dot placement on a digital grid*

The dot pattern is grown alternatively from one sub cell to another. Within each sub cell, the microdot is grown in a clustered spiral fashion. By slicing a large dot into four smaller components, the apparent screen frequency increases by a factor of 2 while the number of gray levels is still high. The drawbacks are the slightly textured contouring and the tone jump in the highlight region that results from alternating growth from nucleus to nucleus.

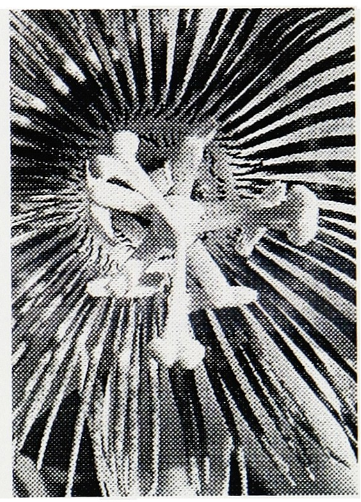
Dispersed Dots. The dispersed dot dither turns binary pixels on or off individually without grouping them into the clusters at a particular tone level in the highlight and the mid-tone regions[11]. It employs fixed size micro dots whose center-to-center spacing (or frequency) is varied according to the density or tone level of the input pixels. For this reason, it is sometimes called "frequency modulation" (FM). In FM screening, the RIP converts the four channels and applies frequency modulated screening to the separated channels. An example of FM screening is shown in Figure 12.



Dither (300dpi)



Continuous Tone



AM (300 dpi, 60 lpi, round dot)

Figure 12. *FM Screening*

The spots are widely spaced in highlight areas and are clustered together in shadow areas of an image. The spacing of spots is determined by the screening algorithm according to the tonal value and the presence of any nearby spots.

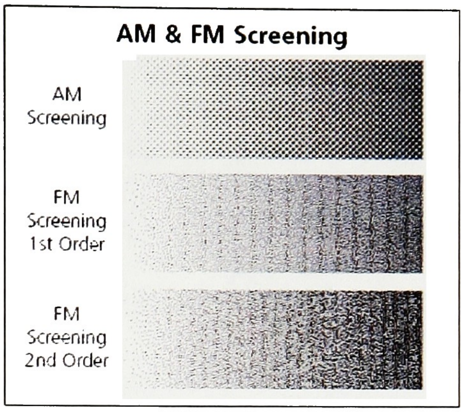


Figure 13. *First order and Second order FM screenings*

There are two types of FM screening methods: first order and second order. A problem of past FM screening has been that some offset presses and proofing systems have had trouble holding this very small stochastic dot. To solve this problem the second generation of FM screening uses a cluster approach which combines very small spots into larger micro-dots.

One advantage that FM screening has over AM screening is that FM screening does not produce moiré and rosette patterns although other patterns, such as dot-cluster formations can occur. When printing more than four colors (e.g. hexachrome), the control of interference patterns in AM screening can be a problem. Many printers using more than four colors choose the FM screening technique for this reason.

Test Targets

Normal graphic arts digital test targets are designed for output devices with high addressability (~1200-3600 spi). The writing method in these devices is binary; ink can only be either on or off. Gray levels are achieved by turning some pixels on or off within a halftone cell. However, there are many output devices which can not resolve very fine pixels; such devices are unable to generate the amount of gray levels and addressability required for AM screening. However, they can still reproduce high quality images, as they are able to apply varying amounts of ink on a given spot, and therefore obtain many gray levels despite their low addressability. For such devices, normal graphic arts digital test targets are unsuitable, as these targets are designed for binary systems with high spatial addressability.

The RIT Contrast Resolution Test Target and the Gutenberg Test Target can both be used on low or high addressability printing systems, providing they have the ability to accept PostScript files. The target's purpose is to measure the relationship between contrast and resolution of a printing system. It is important to note that the target is measuring the capabilities of the printing system and not the printing device.

RIT Contrast Resolution Test Target

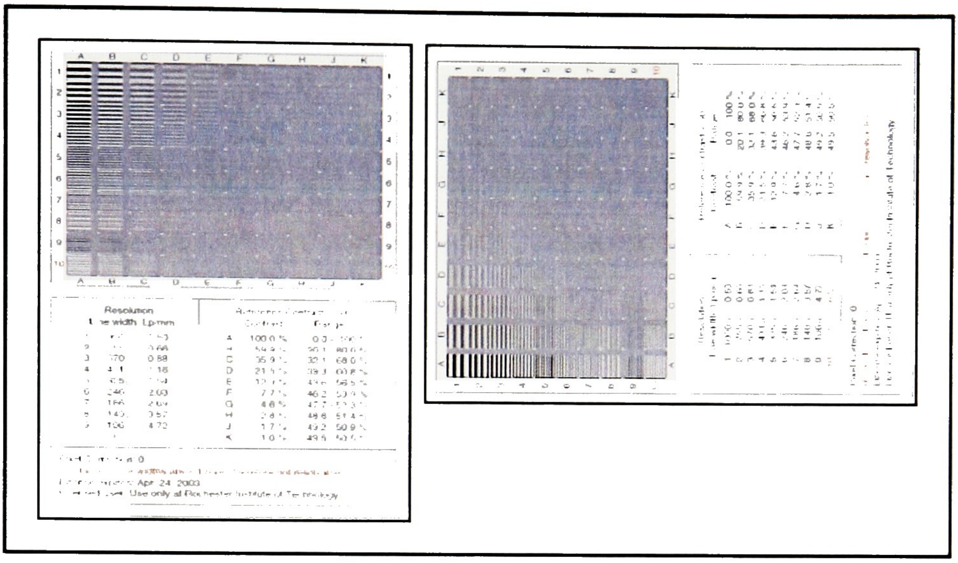


Figure 14. *Black Quadrant Diagram*

The target was developed by Prof. Franz Sigg and Dr. Edward Granger and is shown in Appendix A. On the target, there are two black contrast-resolution quadrants, one horizontal, and the other vertical. The reason for this angular duplication is that many printing systems differ in contrast-resolution capabilities in the horizontal and vertical imaging directions; therefore, the target will indicate such directional contrast-resolution limitations. The target contains pairs of black, magenta, and cyan quadrants; however, for this thesis, only black was used.

Figure 15 above illustrates the quadrant design on the x-imaging axis. The quadrant is constructed as a series of lines along the x-axis. The x-axis is divided into 10 separate "strips" or columns, each separated by small black or white dots across the quadrant. Each strip contains a series of lines with identical line width and spacing. The line width and spacing decreases logarithmically as the strips descend from top to bottom. They start with 1000μm line and space widths, and decrease to 80μm widths.

Each line width is modulated to 10 different contrasts from left to right across the target. The far left strip has maximum contrast applied to the lines (100% tonal value for the colored line, and 0% tonal value for the spacing). The contrasts decrease in logarithmic steps from 100% - 0% (a difference of 100%). to 49.2% - 50.8% (a difference of 1.0 %)[12].

The resolution and contrast ranges of the target can be defined by editing the header in the EPS file.

All strips across the y-axis are centered on a single reference tint value, i.e. the average tonal value between the lines and spacing in each vertical strip remains at a constant throughout the quadrant. By default, this tint value is 50% (mid tone range). However, it may be changed to a reference tint value of 25% (highlight range) or to a reference tint value of 75% (shadow range). This reference value can be changed in the header of the EPS file. There are additional editable fields in the file header such as the line width, the target length, etc., that allow further customization of the target for specific output devices.

Evaluation of the RIT Contrast Resolution Test Target Form

In evaluating the target, each quadrant of the target is visually assessed. First, an observer views the 100% contrast patch at a given x strip, i.e. an observer views the 100% contrast area within the 1000μ boundary. The observer then looks across the selected x-axis strip for the area where he/she can no longer see all the lines which are present in the 100% contrast area. The lines do not have to be perfectly clear but must at least be interpreted as horizontal lines. In the last area where the lines are just visible, the contrast level for that area is recorded. For instance, at the 1000μ strip, if an observer can still distinguish all the horizontal lines that are present in the 100% patch down to 3.9%, and not lower, then 3.9% is the recorded value for the 1000μ x-axis strips in that quadrant. This procedure is then repeated for all nine remaining x-axis strips in that quadrant. If it is determined that the output system was unable to render the lines at a given x-axis strip, then a reading is not required and the whole horizontal strip is ignored[13].

After measurement, the recorded data can be plotted as a contrast sensitivity curve. By taking the two contrast sensitivity curves (quadrants printed both in horizontal and vertical directions), the Contrast-Resolution-Volume (CRV) can be calculated - the methodology for performing the test is explained in Chapter 5. For data analysis purpose, a Microsoft Excel workbook was developed to execute graphing of the contrast sensitivity curves and CRV calculations.

Gutenberg Test Target

The Gutenberg Test Target has been collectively developed by Dr. Edward Granger and Prof. Franz Sigg for measuring resolution (print quality) limitations of an imaging system/device. The Gutenberg Test Target consists of a series of images of Gutenberg, each one with only 24 x 36 pixels. The image of Johannes Gutenberg was chosen for the target as a tribute to his contribution to printing.

The target can be used on both high as well as low addressability systems. Other than measuring the contrast and resolution capability of a printing/imaging system, the target can also act as a visual dot gain indicator because it includes quarter tones, mid-tones and shadows areas. The Gutenberg Test Target form is given in Figure 16.

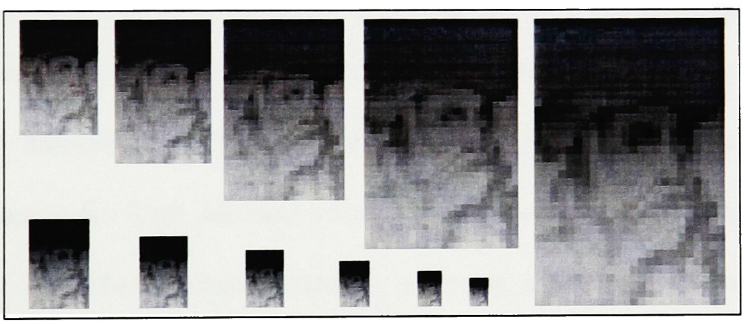


Figure 16. *Gutenberg Test Target*

The Gutenberg Test Target consists of a series of images of Gutenberg, arranged in descending order in size. The size of the Gutenberg images used in the form follows a mathematical series that is provided in Table 1.

0.15	0.22	0.32
0.48	0.71	1.05
1.55	2.29	3.38
5.00	7.39	10.93
16.17		

Table 1. Size *series used for Gutenberg Images*

The target contains eleven Gutenberg images, each image representing the various frequencies (image detail from high to low) that are present in an image.

Evaluation of the Gutenberg Test Target Form

The Gutenberg Test Target is supposed to be evaluated under standard viewing conditions such as a neutral grey background, standard viewing distance and standard ambient light conditions. For evaluation, each image of the Gutenberg Test Target is visually assessed. Starting from the largest Gutenberg image, the observer reads the target horizontally in anti-clockwise direction moving from the largest towards the smallest Gutenberg image in the form. Moving from the largest towards the smallest, the observer indicates (ranks) at which image Gutenberg's face is still recognizable or discernible. Systems with different resolution/contrast capability will yield different results[15].

Once ranked, the results are computed and analyzed.

Endnotes for Chapter 2

1. Sigg Franz. "Definition of Imaging Terms (5th draft)" Rochester, NY: RIT, 2003.
2. Ibid., 5.
3. Ibid., 7.
4. Ibid., 8.
5. http://www.tu-harburg.de/rzt/tuinfo/periph/drucker/Color_Reproduktion/.
5. Ibid., 6.
6. Kang, Henry R. "Digital Color Halftoning" Piscataway, NJ: SPIE, IEEE, 1999.page 213
7. Ibid., 213.
8. Ibid., 252.
9. Ibid., 259.
10. Ibid., 260.
11. Ibid., 279.
12. Harper, Eliot. "An Investigation Into the Relationship Between Contrast and Resolution" Rochester, NY: RIT, 2000.
13. Ibid.
15. Sigg Franz. "Workflow to make Gutenberg target" Rochester, NY: RIT, 2002.

Chapter 3

Review of the Literature

There is relatively limited published work in the area of Subjective Quality Function (SQF) and its correlation with subjective image judgments, although there are several published papers on similar topics. This chapter, first reviews several papers on similar topics, and then identifies the work done by Dr. E. M. Granger in the field of image quality assessment using SQF, and summarizes the findings.

Elliot Harper has published a Master's thesis entitled "An investigation Into the Relationship between Contrast and Resolution of a Printing System Using the RIT Contrast Resolution Test Target.[1] " The hypothesis of this paper was that the RIT Contrast Resolution could provide a method of discriminating the Contrast Resolution Volume (CRV) of marking engines and screening methods by using analysis methods intended for use with the target. The RIT Contrast Resolution Test Target has been developed to measure the relationship between contrast and resolution of a printing system. The target is shown in Appendix A. On the target, there are six contrast-resolution quadrangles; two cyan, two magenta and two black, one for the horizontal (x) imaging axis, and one for the vertical (y) imaging axis.

A graph can be plotted to show resolution as a function of contrast in either the x or y direction. The three axes of contrast resolution volume are resolution in the x direction, resolution in the y direction and contrast in the z direction. In this report, Harper was able to prove that the RIT Contrast Resolution Target could be used to differentiate between the printing systems. He also showed that even though different observers assign different CRV numbers to a given system, there is good agreement between the different observers as to which system is better compared to another system. The correlation of the results from the test with subjective impressions of images was left for further studies.

Yigal Gur analyzed an approach to describe the performance quality of printing systemsen in his paper titled "Image Quality Contrast Transfer and Tone Reproduction " in the 1989 TAGA proceedings. The paper focuses on how the Contrast Transfer Function (CTF) can be used to produce a tone reproduction curve (TRC) for an arbitrary screen frequency. By using such a procedure, a scanner can be adjusted to fit a given ink/paper/press performance. Gur presents a model for producing a TRC from a given CTF through calculating the Modulation Transfer Function (MTF), which allows the image of the known halftone dot to be calculated. The theory was verified by printing a GURLAB target together with a set of 133 lpi circular dot halftone tints on a single color offset lithographic press. From this experiment, it was concluded that two approaches to the description of a half tone system, TRC and CTF, are equivalent.

Yigal Gur and Francis O'Donnel have published a paper entitled "Image Quality Assessment of Ink-Jet Printers" in the 1987 TAGA proceedings. In this paper, Gur and

O'Donnell explain that an ink-jet printer is a binary machine which can either print a dot at a particular location on a sheet of paper or leave it blank. It is the binary nature of this printing process which prevents the use of the classical MTF on an inkjet print. Gur and O'Donnell identify an alternative method for assessing the print quality of an ink-jet printer using a Reflectance Transfer Function (RTF). In their experiment, Gur and O'Donnell obtained the RTF by reflectance measurement of test target containing checkerboards of various frequencies. The RTF is cascaded with the MTF of the human eye to give a subjective quality factor (SQF). The paper notes that this SQF has been shown to correlate well with the subjective assessment of the print quality of an inkjet printer. The paper concludes that this method is effective, and that it provides a simple and cost effective way for instrumental measurements and theoretical analysis of the print quality for an inkjet printer.

Pirrko Oittinen and Hannu Saarelma examined the differing measurements for image quality, which have application in digital reproduction in their paper entitled "Quality in Digital Printing Reproduction" in the 1985 TAGA proceedings. The paper explains that there are several different measurements which can be used to assess quality in digital printing reproduction. These include the bandwidth of the image signal (known as the gamut), fidelity, and information. Oittinen and Saarelma identify that the quality of the images can be measured either as signal properties, or as transfer properties of imaging properties – the former provides a characterization of the gamut of an image, while the latter characterizes the properties of the image transfer system. Transfer type quality measure can either be relative (based on systems theory), or probabilistic (based on information theory), or express the absolute difference of the two images (as distortion type measures do). The paper concludes that both transfer system measures, as well as information-based measures, can be applied as criteria for restorations and enhancement algorithms in digital printing reproduction.

Dr. E. M. Granger in the paper entitled "Image Quality of Digital Cameras" published at X-Rite inc., devises a new method of estimating MTF employing edge target to sample a digital array. In this paper, Granger shows that the edge measurement can then be converted to an average MTF using second moment statistics. The new image quality measure is then easily extended to predict the performance of color CCD cameras. In addition, the second moment can be used to predict the image quality of the displayed image when it is viewed at different magnifications. Granger developed a single parameter "universal" image quality template that allows rapid assessment of system performance.

Dr. E. M. Granger and K. N. Cupery have published a paper entitled "An Optical Merit Function (SQF), which correlates with the subjective image judgments" in 1971 at Eastman Kodak Company. The paper developed a Subjective Quality Factor (SQF) for an objective figure of merit which could be easily calculated and directly measured in practice, and would correlate with subjective rank independent of MTF form. The tests performed included non-symmetric, two-dimensional images and the results showed that SQF was able to predict image quality within normal reader error and was 0.988 correlated with the measured data.

E. M. Granger published a paper entitled "Subjective Assessment and Specification of Color Image Quality" in 1974 in the Society of Photo-Optical Instrumentation Engineers (SPIE). His paper developed a practical image quality criterion which is easily calculated and directly measurable, and which gives consistent evaluations of system performance. In this paper, the image quality merit function is evaluated for a wide variety of MTF shapes which include chromatic and non-symmetrical image errors. The paper proves that the image quality merit function is able to predict image quality within normal reader error and is linearly correlated with the measured data. The tests for the quality criterion for color and black and white images were kept to the physically realizable optical systems producing grain free images and to a range of image quality from excellent to unusable.

Endnotes for Chapter 3

1. Eliot Harper, "*An Investigation Into the Relationship Between Contrast and Resolution*" (Masters Thesis, Rochester, NY: RIT, 2000).
2. Franz Sigg, "*RIT Contrast Resolution Test Target*" (Rochester, NY: RIT, 2000).
3. Yigul Gur, "Image Quality, Contrast Transfer and Tone Reproduction" *1989 TAGA Proceedings*: 470-480.
4. Ibid., 477-478.
5. Ibid., 479.
6. Yigul Gur, and Francis X. D. O'Donnell, "Image Quality Assessment of Ink-Jet Printers" *1987 TAGA Proceedings*: 630-641.
7. Ibid.
8. Ibid.
9. Ibid., 640.
10. Pirkko Oittinen and Hannu Saarelma, "Quality in Digital Printing Reproduction" *1985 TAGA Proceeding*: 621-633.
11. Ibid., 632.
12. Ibid.
13. Edward M. Granger, "Image Quality of Digital Cameras" (Grandville, MI: X-Rite, Inc., 1972).
14. Ibid.
15. Edward M. Granger and K. N. Cupery "An Optical Merit Function (SQF), Which Correlates with Subjective Image Judgments" *1972 SPSE Proceedings*: 221-230.
16. Ibid., 229.
17. Edward M. Granger. "Image Assessment & Specifications" *1974 SPIE Proceedings, Volume 46*: 86-92.
18. Ibid., 92.

Chapter 4

Hypotheses

H_0: The Gutenberg Test Target cannot discriminate between the resolution capabilities of different printing systems.

H_1: The Gutenberg Test Target can discriminate between the resolution capabilities of different printing systems.

H'_0: The Gutenberg Test Target does not provide an easier method to evaluate print quality than the RIT Contrast Resolution Target.

H'_1: The Gutenberg Test Target can provide an easier method to evaluate print quality than the RIT Contrast Resolution Target.

Chapter 5

Methodology

The objective of this thesis is to verify that the Gutenberg Test Target can discriminate between different printing systems and provide an easier method to evaluate print quality than the RIT Contrast Resolution Target. To compare their performances, a visual experiment was performed on each target and the results were analyzed and compared.

In order to test these targets, we could have printed images on different output devices and this could have provided us with targets printed with different levels of quality. Instead we chose to simulate different quality levels by blurring the target images to different degrees in Photoshop, and then printing all of them on a single high quality output device. The procedure for performing the experiment is explained in this chapter.

RIT Contrast Resolution Test Target

Construction of Target

Using Photoshop 6, a test form was created with the RIT Contrast Resolution target placed in the x and y direction, as shown in Figure 16.

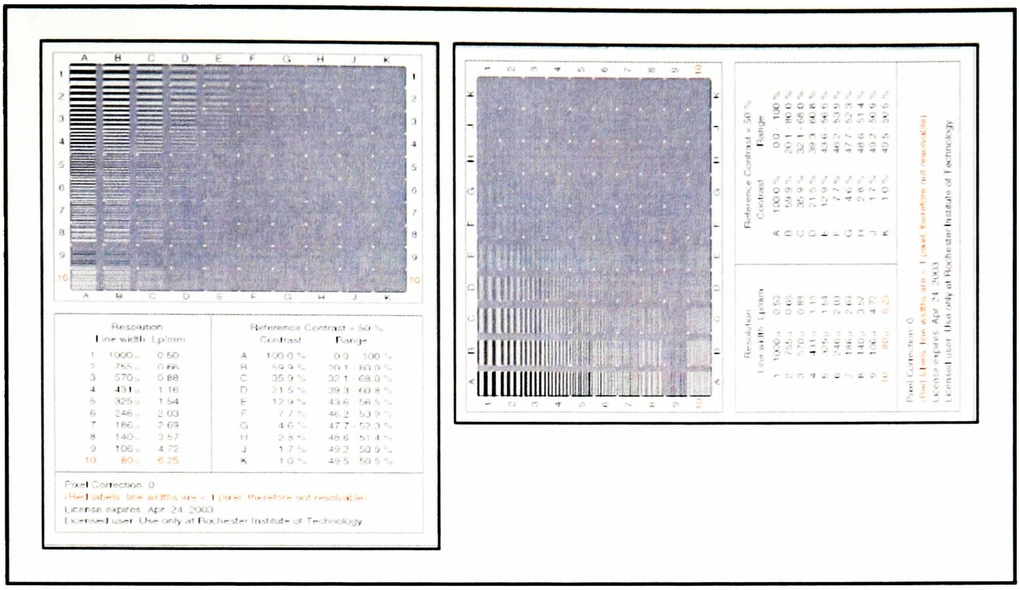


Figure 16. *RIT Contrast Resolution Test Target form*

Fourteen such test targets were created and each one was blurred to a different degree using the Gaussian blur function in Photoshop 6. The series of Gaussian blur values applied to the test target are given in Table 1. Each value was multiplied by 1.5 to obtain the next higher value, resulting in a logarithmic scale. The resulting spacing is not necessarily perceptually equidistant.

The saved eps test target images were then placed in QuarkXPress and sent to output on Polaroid PolaProof. The device was chosen because of its high addressability (2400 spots/inch), i.e. its capability to produce a high number of gray levels at high printing resolutions. Each target was labeled on the back with the corresponding Gaussian blur value associated with it.

Target No.	Gaussian Blur Value	Target No.	Gaussian Blur Value	Target No.	Gaussian Blur Value
Target 1	0	Target 6	0.71	Target 11	5.0
Target 2	0.15	Target 7	1.05	Target 12	7.4
Target 3	0.22	Target 8	1.55	Target 13	11.0
Target 4	0.32	Target 9	2.3	Target 14	16.2
Target 5	0.48	Target10	3.4		

Table 2. *Gaussian Blur values used for the targets*

Evaluation of targets

Before target evaluation, each observer was instructed in target composition, and the process of visually recording the x strips on the target was explained. Next, each observer was presented with test targets to verify that they understood the visual assessment process of the targets. Once it was determined that the observers understood the evaluation process, the targets were placed in a randomized order to remove any kind of subjectivity and were presented to the observer under standardized conditions of D50 illumination, neutral gray board background and a 2x power stand magnifier for viewing targets. Nine observers were used for the visual assessment. Each of the observers was presented with the fourteen RIT Contrast Resolution Test Targets printed on the Polaroid Polaproof, and their responses were recorded.

From the recorded data, a contrast sensitivity curve, as shown in Figure 17, was plotted for each blur value for each observer. The y-axis on the graph indicates measured contrast level and the x-axis on the graph indicates resolution. The graph's excel worksheet used for analysis was constructed in such a way that the user could choose the units for the x-axis, the actual line width, in either cycles/mm or cycles/in.

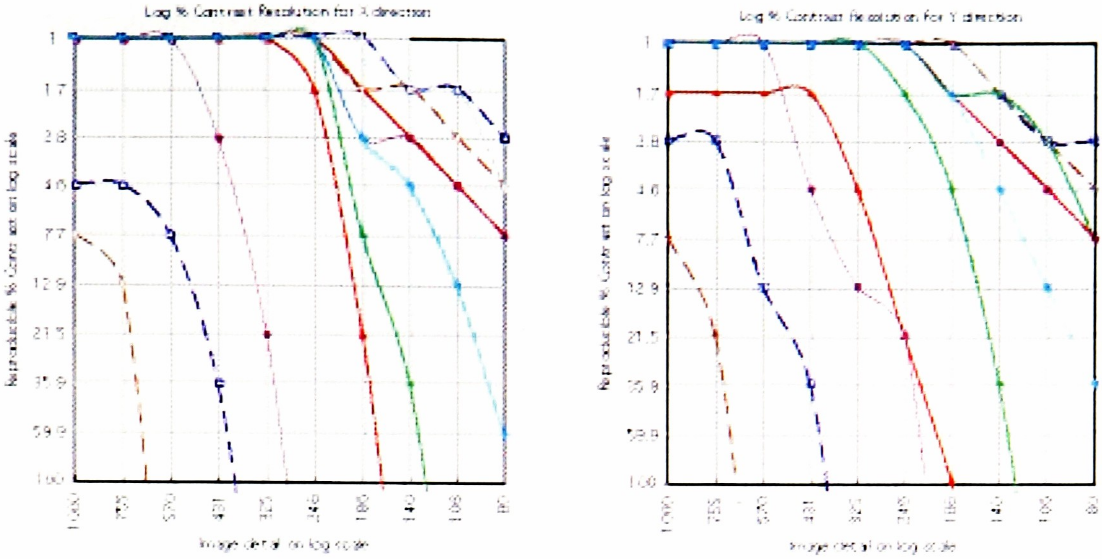


Figure 17. *Contrast sensitivity curves and relative CRV values in x direction for Observer1*

Once the CS curves for the horizontal and vertical printed directions were plotted for a single color on the target, the CRV was calculated. CRV can be calculated by different methods. To calculate the volume, three dimensions were required: resolution in the x direction, resolution in the y direction, and contrast. To keep things simple, the following method was used. Each step on the axis was taken as one unit for that axis. There are ten steps on each log scale, which means that the highest possible volume that could be achieved is 1000. The quantization is relatively coarse as only 10 steps per dimension are used. The CRV numbers are relative to the arbitrarily chosen range of resolution and contrast; therefore, it is referred to as relative volume. The CRV formula is:

$$\sum_{\text{Contrast step 1}}^{\text{Contrast step 10}} \left(\begin{array}{l}\text{Number of}\\ \text{log steps in}\\ \text{x direction}\\ \text{for a given}\\ \text{contrast}\end{array}\right) * \left(\begin{array}{l}\text{Number of}\\ \text{log steps in}\\ \text{y direction}\\ \text{for a given}\\ \text{contrast}\end{array}\right)$$

A three dimensional plot of CRV was constructed using MS Excel; it was only used for visual representation purposes, and not for calculation. This is shown in figure 3.

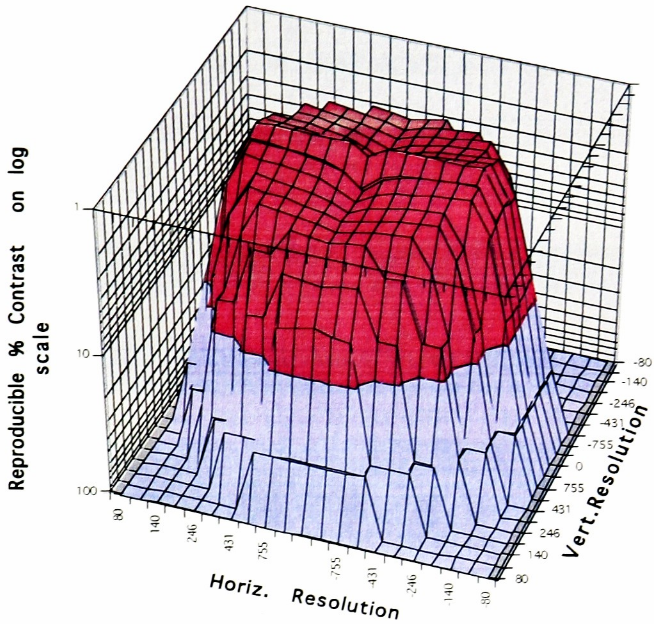


Figure 18. *A three-dimensional distribution of contrast-resolution-volume*

Gutenberg Test Target

Construction of Targets

With the help of the Gaussian Blur function in Photoshop 6, fourteen Gutenberg test targets were created by assigning the same series of blur values (Table 1) as was used with the RIT Contrast Resolution Test Target. The target is shown in Figure 19.

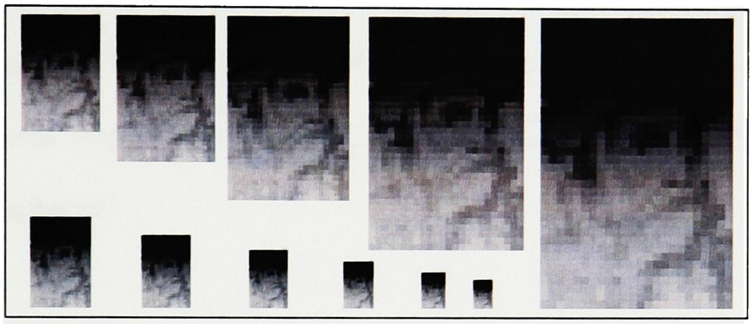


Figure 19. *Gutenberg Test Target form*

The test targets were saved as an eps, placed in a QuarkXPress document and were sent to output on Polaroid PolaProof. This device was chosen because of its high addressability (2400 spots/inch), i.e. its capability to produce a high number of gray levels at high printing resolutions. For this test, PolaProof was set to 2400 dpi at 150-line screen with a Euclidean dot shape. The proof was output onto Krome Kote paper using SWOP donor sheets with a semi-matte finish material.

Evaluation of targets

After printing, the targets were cut out and labeled on the back with the corresponding Gaussian blur values. Before the target evaluation, each observer was instructed in target composition, and the process of visually ranking the targets was explained. Next, each observer was presented with a series of targets to verify that the observer understood the ranking process of the target forms, i.e. the Gutenberg image with the least visually acceptable quality. Once it was determined that the observers understood the evaluation process of the target, the targets were placed in a randomized order to remove any kind of subjectivity and presented to the observers under standardized conditions, i.e. neutral grey background and viewed under a 2x power stand magnifier with D50 ambient light conditions. Nine observers were used for the visual assessment. Each one was presented with the fourteen targets and their rankings were recorded. Next, the results were computed and analyzed.

Chapter 6

Results

RIT Contrast Resolution Test Target

In order to test the hypothesis, the averaged CRV's for each Gaussian blur value were compiled in Table 3. This data was analyzed by a two way ANOVA test where the two dimensions were the fourteen levels of Gaussian blur and the nine observers. The calculated ANOVA data is given in Table 4.

#	Reader 1	Reader 2	Reader 3	Reader 4	Reader 5	Reader 6	Reader 7	Reader 8	Reader 9	AVG	STDEV	+2STDEV	-2STDEV
1	894.0	886.0	909.0	730.0	824.0	741.0	864.0	553.0	714.0	790.56	116.60	1023.8	557.35
2	845.0	755.0	894.0	680.0	802.0	761.0	779.0	581.0	718.0	757.22	91.83	940.9	573.55
3	830.0	830.0	870.0	661.0	877.0	760.0	833.0	604.0	757.0	780.22	94.46	969.1	591.30
4	823.0	751.0	857.0	680.0	894.0	720.0	791.0	530.0	729.0	752.78	108.26	969.3	536.25
5	921.0	799.0	928.0	646.0	935.0	776.0	827.0	489.5	800.0	791.28	145.77	1082.8	499.73
6	887.0	798.0	895.0	630.0	891.0	774.0	750.0	628.0	746.0	777.67	103.06	983.8	571.55
7	902.0	742.0	902.0	644.0	838.0	739.0	682.0	503.0	725.0	741.89	127.98	997.9	485.92
8	707.0	661.0	751.0	605.0	788.0	753.0	509.0	509.0	676.0	662.11	102.69	867.5	456.73
9	600.0	589.0	756.0	494.0	747.0	711.0	697.0	416.0	561.0	619.00	117.88	854.8	383.25
10	483.0	532.0	734.0	378.0	640.0	440.0	511.0	400.0	414.5	503.61	118.04	739.7	267.53
11	325.5	416.0	569.0	245.5	441.0	417.0	427.0	184.0	265.5	365.61	119.92	605.4	125.78
12	202.0	150.0	234.0	140.0	236.0	123.0	196.0	91.0	137.0	167.67	51.16	270.0	65.34
13	78.0	62.0	146.0	62.0	92.5	72.0	158.0	44.0	34.5	83.22	42.70	168.6	-2.17
14	19.0	28.0	22.0	21.0	86.0	32.0	27.0	16.5	14.5	29.56	21.91	73.4	-14.26
AVG	608.3	571.4	676.2	472.6	649.4	558.5	575.1	396.4	520.9				
STDEV	315.5	284.2	300.7	244.0	294.2	275.9	268.2	209.3	278.8				
+2STDEV	1239.3	1139.7	1277.5	960.7	1237.8	1110.4	1111.5	814.9	1078.4				
-2STDEV	-22.7	3.0	74.9	-15.4	61.0	6.6	38.7	-22.2	-36.7				

Table 3. *Relative CRV measurements for each observer for each Gaussian blur value for the RIT Contrast Resolution Test Target*

To test whether there was any variance in the average CRV value for each blur value, the null and the alternate hypotheses used in the ANOVA tests were

$H_0: = \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu_8 = \mu_9 = \mu_{10} = \mu_{11} = \mu_{12} = \mu_{13} = \mu_{14}$

At least one population mean is different.

Where μ_{num} = (Average CRV value)

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Gaussian Blur	9262182.59	13.00	712475.58	213.51	0.00	2.31
Observers	841744.46	8.00	105218.06	31.53	0.00	2.69
Error	347044.82	104.00	3336.97			
Total	10450971.87	125.00				

Table 4. *Two-way ANOVA without replication using an alpha level of 0.01 for RIT Contrast Resolution Test Target*

The calculated ANOVA data is given in Appendix C. Using an alpha level of 0.01, $df_{system} = 13$ and $df_{res.error} = 104$, an *F* ratio of 213.51 was obtained, which is much larger than the critical value, proving that the RIT Contrast Resolution Test Target does provide a method of discriminating the CRV of marking engines and screening methods. H_1 is accepted.

Furthermore, the ANOVA test shows whether each observer produced different results.

Using an alpha level of 0.01, $df_{system} = 8$ and $df_{res.error} = 104$, the results show the F ratio for "Observers" is 31.53. Therefore, as the F ratio > 2.68, it is proved that each observer produced different results. Each observer had a different mean for different Gaussian blur values. Hence, H_1 accepted.

To examine the variance among the observer's individual results, first the CRV values were sorted in ascending order for systems, and then by the observers. Next, the mean and standard deviations were calculated for each observer and blur values. This data is provided in Table 5.

#	Reader 9	Reader 8	Reader 1	Reader 4	Reader 3	Reader 7	Reader 2	Reader 6	Reader 5	AVG	STDEV	+2STDEV	-2STDEV
14	14.5	16.5	19.0	21.0	22.0	27.0	28.0	32.0	86.0	29.56	21.91	73.4	-14.26
13	34.5	44.0	78.0	62.0	146.0	158.0	62.0	72.0	92.5	83.22	42.70	168.6	-2.17
12	137.0	91.0	202.0	140.0	234.0	196.0	150.0	123.0	236.0	167.67	51.16	270.0	65.34
11	265.5	184.0	325.5	245.5	569.0	427.0	416.0	417.0	441.0	365.61	119.92	605.4	125.78
10	414.5	400.0	483.0	378.0	734.0	511.0	532.0	440.0	640.0	503.61	118.04	739.7	267.53
9	561.0	416.0	600.0	494.0	756.0	697.0	589.0	711.0	747.0	619.00	117.88	854.8	383.25
8	676.0	509.0	707.0	605.0	751.0	509.0	661.0	753.0	788.0	662.11	102.69	867.5	456.73
4	729.0	530.0	823.0	680.0	857.0	791.0	751.0	720.0	894.0	752.78	108.26	969.3	536.25
3	757.0	604.0	830.0	661.0	870.0	833.0	830.0	760.0	877.0	780.22	94.46	969.1	591.30
2	718.0	581.0	845.0	680.0	894.0	779.0	755.0	761.0	802.0	757.22	91.83	940.9	573.55
6	746.0	628.0	887.0	630.0	895.0	750.0	798.0	774.0	891.0	777.67	103.06	983.8	571.55
1	714.0	553.0	894.0	730.0	909.0	864.0	886.0	741.0	824.0	790.56	116.60	1023.8	557.35
7	725.0	503.0	902.0	644.0	902.0	682.0	742.0	739.0	838.0	741.89	127.98	997.9	485.92
5	800.0	489.5	921.0	646.0	928.0	827.0	799.0	776.0	935.0	791.28	145.77	1082.8	499.73
AVG	520.86	396.36	608.32	472.61	676.21	575.07	571.36	558.50	649.39				
STDEV	278.78	209.28	315.51	244.02	300.66	268.20	284.16	275.94	294.18				
+2STDEV	1078.41	814.92	1239.35	960.66	1277.54	1111.46	1139.68	1110.38	1237.76				
-2STDEV	-36.70	-22.21	-22.70	-15.44	74.89	38.68	3.04	6.62	61.02				

Table 5. *Relative CRV measurements for each observer for each Gaussian blur in ascending order for blur and observers for RIT Contrast Resolution Test Target*

From this sorted data, a graph for the average CRV responses of the observer versus the blur values was plotted. This graph is shown in Figure 20.

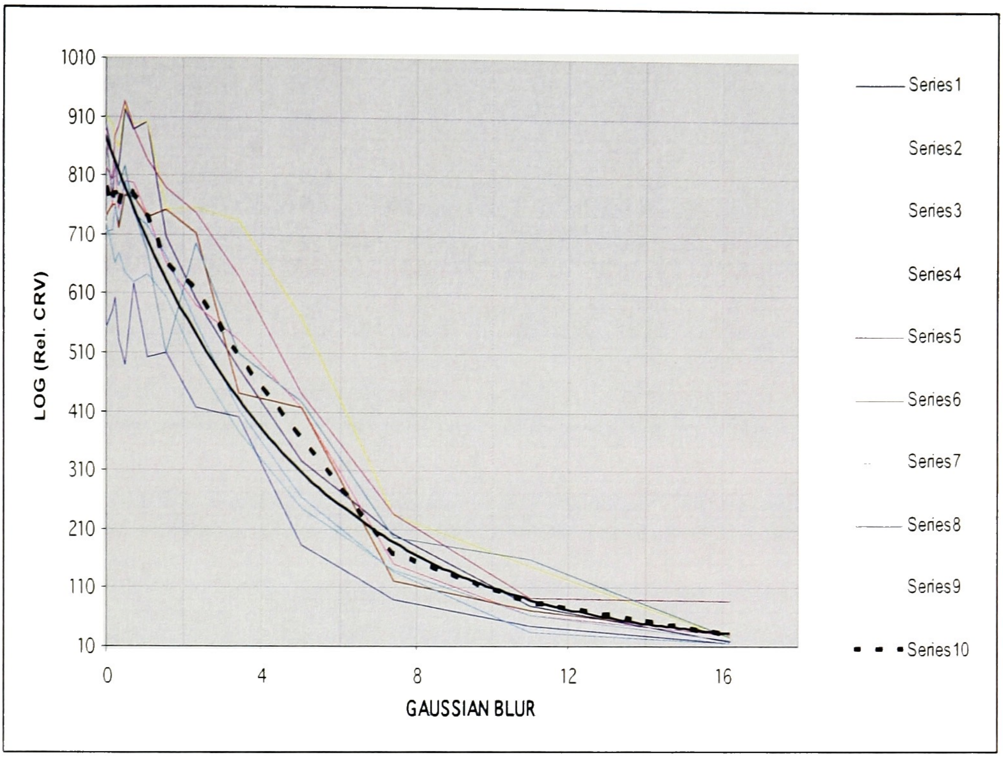


Figure 20. *Blur values versus CRV values for RIT Contrast Resolution Test Target*

From the curve distribution for each observer's data, it can be noted that the observers differed in evaluating the target. However, each observer's data fell within a certain response range relative to the average of all distribution. Although the judging criteria was different for each observer, each observer evaluated the different systems using a reasonably consistent assessment criteria. This trend was analyzed further by calculating the normal distribution percentile for each CRV value. This data is given in Table 6. These percentile values were then plotted against their corresponding CRV values as a normality test. This graph is shown in Figure 21.

#	Reader 8	Reader 9	Reader 4	Reader 6	Reader 5	Reader 7	Reader 2	Reader 1	Reader 3
5	1.9%	52.4%	15.9%	45.8%	83.8%	59.7%	52.1%	81.3%	82.6%
4	2.0%	41.3%	25.1%	38.1%	90.4%	63.8%	49.3%	74.2%	83.2%
1	2.1%	25.6%	30.2%	33.5%	61.3%	73.6%	79.3%	81.3%	84.5%
2	2.7%	33.5%	20.0%	51.6%	68.7%	59.4%	49.0%	83.0%	93.2%
7	3.1%	44.8%	22.2%	49.1%	77.4%	32.0%	50.0%	89.5%	89.5%
3	3.1%	40.3%	10.3%	41.5%	84.7%	71.2%	70.1%	70.1%	82.9%
9	4.3%	31.1%	14.4%	78.2%	86.1%	74.6%	40.0%	43.6%	87.7%
11	6.5%	20.2%	15.8%	66.6%	73.5%	69.6%	66.3%	36.9%	95.5%
12	6.7%	27.4%	29.4%	19.1%	90.9%	71.0%	36.5%	74.9%	90.3%
8	6.8%	55.4%	28.9%	81.2%	89.0%	6.8%	49.6%	66.9%	80.7%
6	7.3%	37.9%	7.6%	48.6%	86.4%	39.4%	57.8%	85.6%	87.3%
14	16.0%	9.3%	40.2%	94.2%		77.2%	81.9%	28.1%	46.7%
13	17.9%	12.7%	31.0%	39.6%	58.6%	96.0%	31.0%	45.1%	92.9%
10	19.0%	22.5%	14.4%	29.5%	87.6%	52.5%	59.5%	43.1%	97.5%

Table 6. *Normal distribution percentile for each observer on each system for RIT Contrast Resolution Test Target*

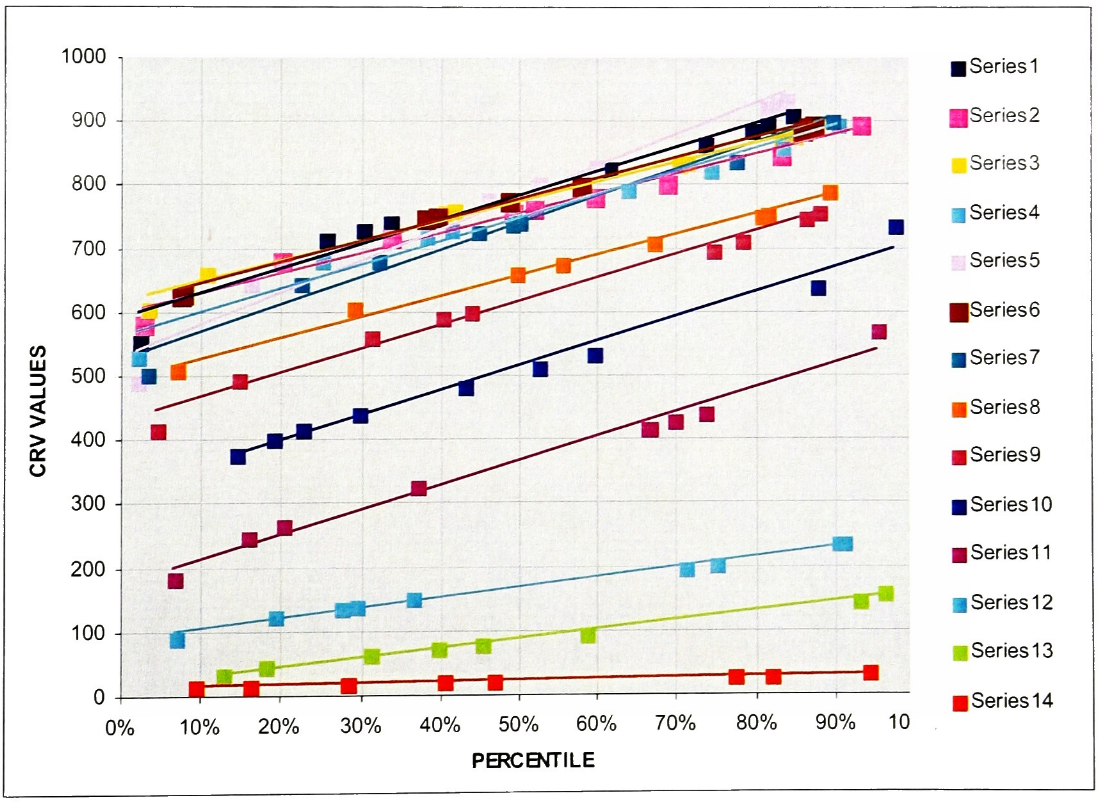


Figure 21. *Normality test of observers for RIT Contrast Resolution Test Target*

Linear trend-lines have been added to the graph to indicate the distribution of the observer data for each system. The graphs indicate that almost all of the observers' data falls close to these lines for each system, and all the lines fall almost parallel to each other. As the data from each observer falls almost on a straight line, it can be concluded that the data for each system is normally distributed. This is significant, as it shows that all observers had a similar deviation about the mean (trend-lines fall parallel to each other), and it proves that each observer used consistent criteria to evaluate the different Gaussian blur values. Furthermore, as many of the trend-lines are separated, it shows that the observers ranked the system in the same relative order. Although the target allows observers to differentiate between printing systems and rank systems in almost the same sequence, the reading of the target is subjective. Subjectivity in an observer is an unavoidable factor of noise. It can be minimized through training, but not eliminated.

Gutenberg Test Target

The averaged rankings for each Gaussian blur value were compiled into a Table 7. This data was entered into a two-way ANOVA test, where the two dimensions in the system were the Gaussian blur and the observer's rankings. The calculated ANOVA data is given in Table 6.

#	Obs 1	Obs 2	Obs 3	Obs 4	Obs 5	Obs 6	Obs 7	Obs 8	Obs 9	AVG	STDEV
1	10	11	11	10	10	11	11	11	10	10.6	0.5
2	10	11	11	11	9	10	11	11	9	10.3	0.9
3	10	11	11	10	9	10	11	11	9	10.2	0.8
4	10	11	10	10	9	11	11	11	8	10.1	1.1
5	10	10	11	11	9	10	11	10	8	10.0	1.0
6	9	11	9	9	9	8	11	10	9	9.4	1.0
7	9	11	9	11	9	9	10	8	8	9.3	1.1
8	8	10	9	10	7	8	9	8	6	8.3	1.3
9	7	8	8	8	7	8	9	7	6	7.6	0.9
10	6	9	6	5	5	6	8	4	4	5.9	1.7
11	5	4	4	5	5	4	6	3	3	4.3	1.0
12	4	5	2	2	3	3	5	2	2	3.1	1.3
13	1	1	1	1	1	1	2	1	0	1.0	0.5
14	0	0	0	0	0	0	0	0	0	0.0	0.0
AVG	7.1	8.1	7.3	7.4	6.6	7.1	8.2	6.9	5.9		
STDEV	3.4	3.9	4.0	4.0	3.3	3.7	3.6	4.1	3.5		

Table 7. *Ranking for each observer for each Gaussian blur value for Gutenberg Test Target*

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	1576.61	13	121.28	209.11	3.1E-68	2.31
Columns	56.57	8	7.07	12.19	3.6E-12	2.69
Error	60.32	104	0.58			
Total	1693.50	125				

Table 8. *Two-way ANOVA without replication using an alpha level of 0.01 for Gutenberg Target*

To test whether there was any variance in the average ranking by the observers for each blur value, the null, and the alternate hypothesis used in the ANOVA tests were

H_0: $= \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_7 + \mu_8 + \mu_9 + \mu_{10} + \mu_{11} + \mu_{12} + \mu_{13} + \mu_{14}$

H_1: At least one population mean is different.

Where μ_{num} = Ranking mean for the blur value

The calculated ANOVA data is given in Appendix C. Using an alpha level of 0.01, $df_{system} = 13$ and $df_{res.error} = 104$, the decision rule for this test is 'Reject H_0 and accept H_1 if calculated F ratio is > 2.88. Otherwise, accept H0'. The results of the ANOVA test are given in Table 3. The *F* ratio for "blur values" is 209.10. Therefore, as the F ratio >2.88, H_0 is rejected and H_1 is accepted. This test proves H_1 of the hypothesis for this thesis that the Gutenberg Test Target provides a method of discriminating "apparent image quality" reproduced by devices using different marking engines and screening methods. Furthermore, to test whether each observer produced a different result, the null and the alternate hypothesis used in the ANOVA tests were:

H_0: $= \mu_1 + \mu_2 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_7 + \mu_8 + \mu_9$

H_1: At least one population mean is different

(Where μ_{num} = Ranking mean for the observer)

Using an alpha level of 0.01, $df_{system} = 8$ and $df_{res.error} = 104$, the results show the F ratio for "Observers" is 31.53. Therefore, as the F ratio > 2.68, it is proved that each observer produced different results. Each observer had a different mean for different Gaussian blur values.

To examine the variance between the observers' individual results, first the CRV values were sorted in ascending order for systems, and then by the observers. Next, the mean and standard deviations were calculated for each observer and blur value. This data is provided in Table 9.

#	OBS 9	OBS 5	OBS 8	OBS 1	OBS 6	OBS 3	OBS 4	OBS 2	OBS 7	AVG	STDEV
14	0	0	0	0	0	0	0	0	0	0.0	0.0
13	0	1	1	1	1	1	1	1	2	1.0	0.5
12	2	3	2	4	3	2	2	5	5	3.1	1.3
11	3	5	3	5	4	4	5	4	6	4.3	1.0
10	4	5	4	6	6	6	5	9	8	5.9	1.7
9	6	7	7	7	8	8	8	8	9	7.6	0.9
8	6	7	8	8	8	9	10	10	9	8.3	1.3
7	8	9	8	9	9	9	11	10	10	9.2	1.0
6	9	9	10	9	8	9	9	11	10	9.3	0.9
5	8	9	10	10	10	11	9	11	10	9.8	1.0
4	8	9	11	10	11	10	10	11	11	10.1	1.1
2	9	9	11	10	10	11	10	11	11	10.2	0.8
3	9	9	11	10	10	11	11	11	11	10.3	0.9
1	10	10	11	10	11	11	10	11	11	10.6	0.5
AVG	5.9	6.6	6.9	7.1	7.1	7.3	7.2	8.1	8.1		
STDEV	3.5	3.3	4.1	3.4	3.7	4.0	3.9	3.9	3.5		

Table 9. *Relative Rankings for each observer for each Gaussian blur values in ascending order for blur and observers for Gutenberg Test Target*

From this sorted data, a graph for ranking responses of the observer versus the blur values was plotted. The graph is shown in Figure 22.

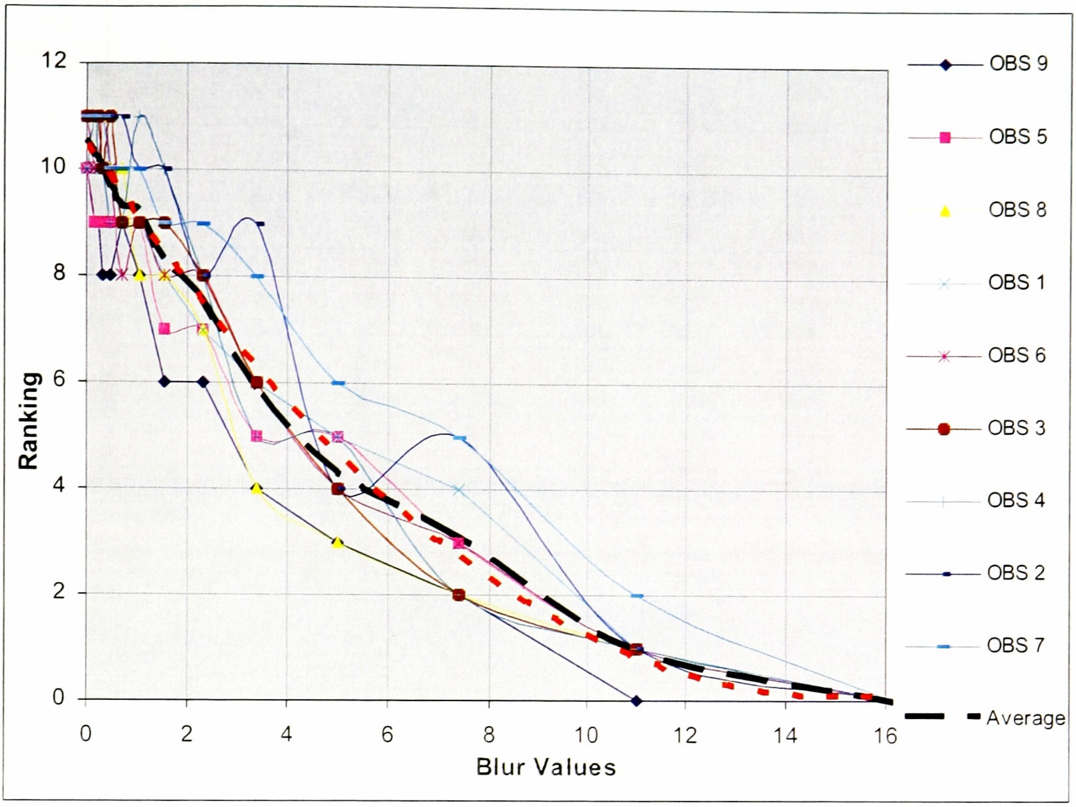


Figure 22. *Blur versus Ranking for Gutenberg Test Target*

From the curve distribution for each observer's data, it can be noted that the observers differed in evaluating the target. However, each observer's data fell within a certain response range relative to the average of all distribution. Although the judging criterion was different for all observers, each observer evaluated the different systems using reasonably consistent assessment criteria. This trend was analyzed further by calculating a normal distribution percentile for each CRV value. This data is shown in Table 10.

#	OBS 9	OBS 5	OBS 8	OBS 1	OBS 6	OBS 3	OBS 4	OBS 2	OBS 7
16.2									
11	2%	50%	50%	50%	50%	50%	50%	50%	98%
7.4	19%	47%	19%	76%	47%	19%	19%	93%	93%
5	9%	75%	9%	75%	37%	37%	75%	37%	95%
3.4	13%	30%	13%	53%	53%	53%	30%	97%	89%
2.3	4%	26%	26%	26%	69%	69%	69%	69%	95%
1.55	4%	16%	40%	40%	40%	69%	90%	90%	69%
1.05	10%	41%	10%	41%	41%	41%	97%	79%	79%
0.71	35%	35%	78%	35%	6%	35%	35%	97%	78%
0.48	3%	21%	59%	59%	59%	90%	21%	90%	59%
0.32	2%	15%	80%	46%	80%	46%	46%	80%	80%
0.22	7%	7%	82%	39%	39%	82%	39%	82%	82%
0.16	6%	6%	78%	35%	35%	78%	78%	78%	78%
0	15%	15%	80%	15%	80%	80%	15%	80%	80%

Table 10. *Normal distribution percentile for each observer for each blur value for the Gutenberg Test Target*

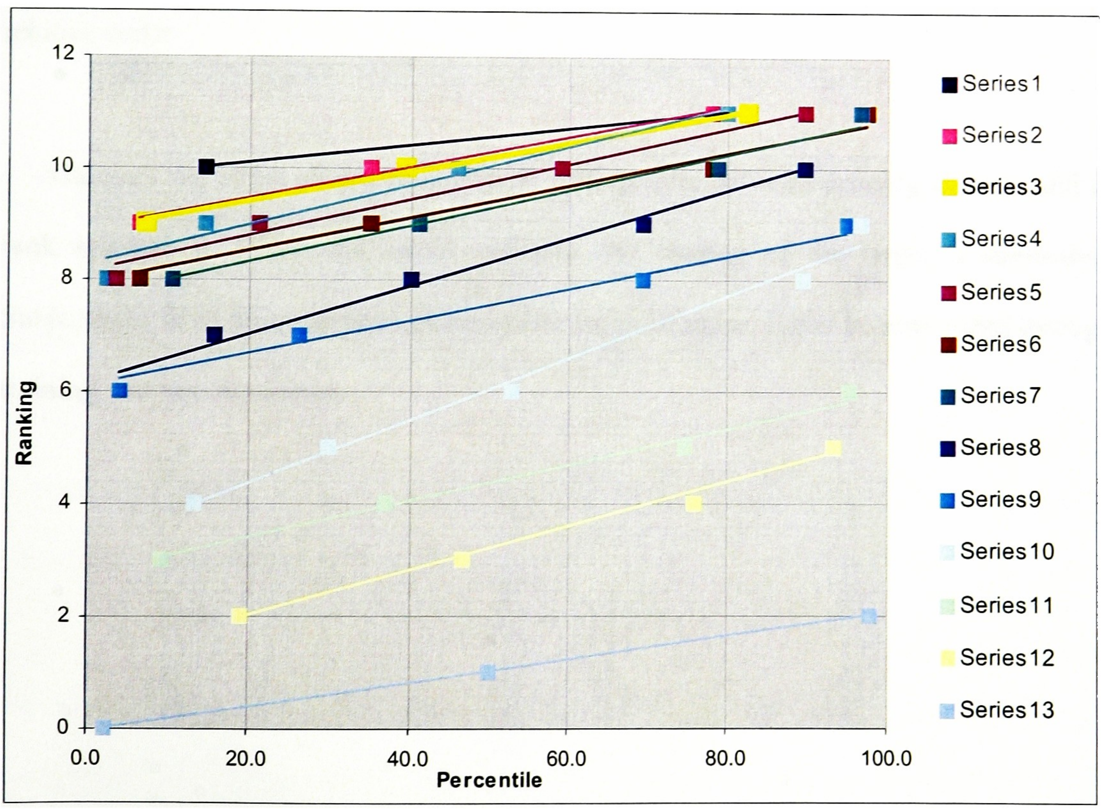


Figure 23. *Normality test of observers for the Gutenberg Test Target Ranking*

These percentile values were then plotted against their corresponding CRV values as a normality test. This graph is shown in Figure 23. Linear trend-lines have been added to the graph to indicate the distribution of the observer data for each system. The graphs indicate that almost all of the observers' data falls close to these lines for each system, and all the lines fall almost parallel to each another. As the data from each observer falls almost on a straight line, it can be concluded that the data for each system is normally distributed. This is significant, as it shows that all observers had a similar deviation about the mean (trend-lines fall parallel to each other), and it proves that each observer used consistent criteria to evaluate the different Gaussian blur values. Furthermore, as many of the trend-lines are separated, it shows that the observers ranked the system in the same relative order.

Although the target allows observers to differentiate between printing systems and to rank systems in almost the same sequence, the reading of the target is subjective. Subjectivity in an observer is an unavoidable factor of noise. It can be minimized through training, but not eliminated.

Chapter 7

Summary and Conclusion

Summary

Different systems differ in both contrast and resolution capabilities, and while a device may have a high spatial addressability, the ability of such a device to render low levels of contrast at a given resolution may be poor. Such limitations of contrast and resolution of an imaging/printing system could be attributed to a combination of various factors such as the screening methods used by the RIP, the image transfer methods of the output device, the substrate used, and the PostScript interpreter.

Both targets, the RIT Contrast Resolution Test Target and the Gutenberg Test Target, are used to determine contrast resolution capability/limitation of a printing system. These targets can be used by any system, which can process a Postscript or an Encapsulated PostScript file. However, the targets differ from each other in two aspects: methodology and the utility of the target.

For the target evaluation, both targets employ different methodologies. The RIT Contrast Resolution Test Target has a complicated construction, and employs a somewhat lengthy procedure using both subjective and objective methods to calculate CRV. The Gutenberg Test Target on the other hand has a simpler construction and only uses visual assessment of the target to give a measure of print quality.

The RIT Contrast Resolution Test Target provides a composite CRV value for the contrast and resolution limit, and also visually displays the contrast resolution curve for each C, M, and K color for the printing /imaging system. Alternatively, the Gutenberg Test Target provides a purely subjective method for print quality assessment and can also be used as a visual indicator of dot gain and gray balance because it contains quarter tones, mid tones and shadows. Similarly, the RIT Contrast Resolution Target can be also set to different reference tonal values (25%, 50%, and 75%) to evaluate contrast resolution capability for different tone reproductions and hence, could be used to test the contrast resolution relationship at different tone value levels. Only the 50% level was tested in this experiment.

The two-way ANOVA test performed on the results obtained from both the targets proved that both targets can discriminate between various printing/imaging devices and systems having different contrast resolution capabilities. The test also shows that the observers were using different criteria for evaluating systems. The analysis of data from both targets showed that although observers appear to use different criteria, they do so in a consistent manner.

When the R^2 (coefficient of determination) values were computed for the average CRV responses versus the blur values for the RIT Contrast Resolution Test Target and the ranking average versus blur values for the Gutenberg Test Target, the RIT Contrast

Resolution Test Target provided a value of 0.991 and the Gutenberg Test Target 0.996. The slight difference in the values is considered insignificant. Furthermore we performed regression on the data obtained by visual experiments performed on the RIT Contrast Resolution Target and the RIT Gutenberg Test Target. The regression preformed on the data provided a correlation of 0.982, and the relation between the two targets came out to be Gutenberg = CRV * 0.128. Figure 24 shows the relationship between Gaussian blur values versus the average CRV values and the average Gutenberg Test Target rankings.

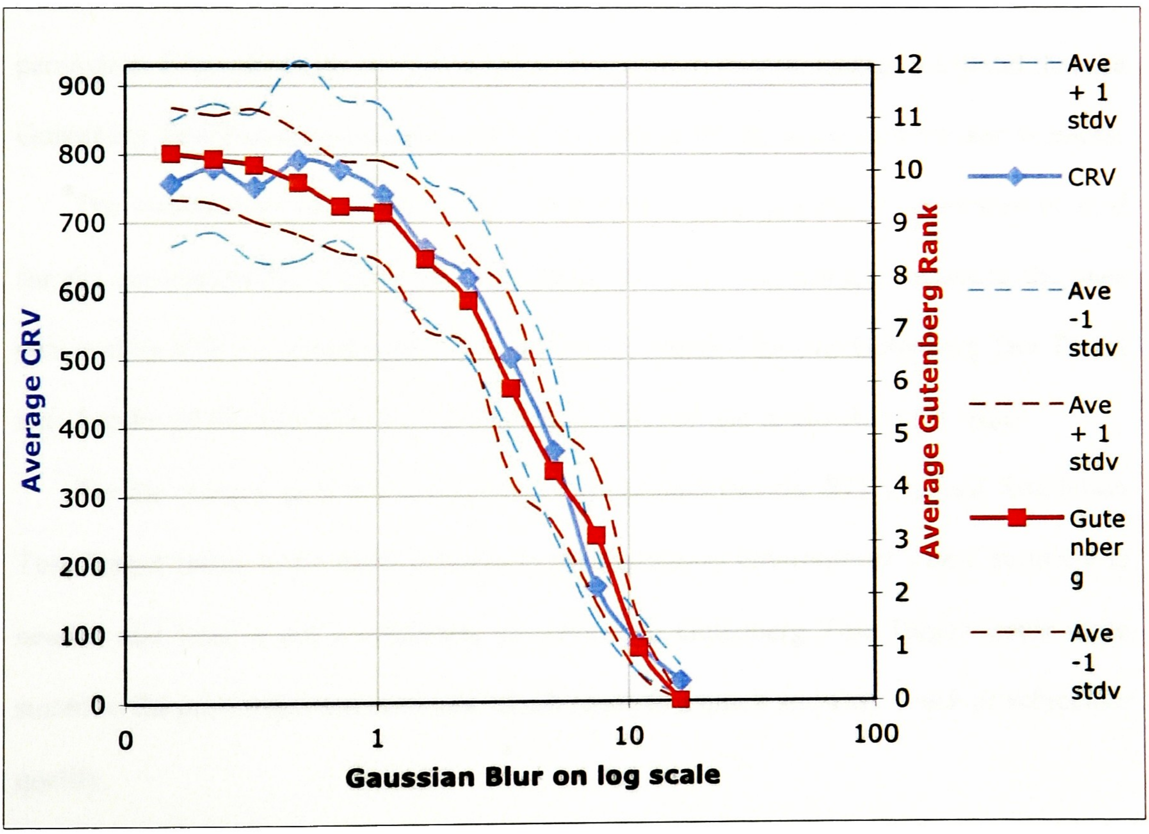


Figure 24. *Relationship between the Gaussian blur values versus the average CRV and the average rankings.*

Therefore, from the results obtained it can be deduced that for the given range of the data for this experiment (small population size, 9 observers) the two methods give very similar results.

It was found out that although both targets perform the same job and with same relative accuracy, the RIT Contrast Resolution Test Target presents a more comprehensive picture of the process by providing a comprehensive multidimensional value of CRV for the contrast resolution capability of a system. In contrast the Gutenberg Test Target provides a very subjective visual ranking giving a single dimension quality parameter. From the point of view of the ease of using the test targets we found that the Gutenberg Test Target was simpler to use as it was relatively easier and quicker to assess.

The results from our analysis show that for the small population of observers used for the experiment the Gutenberg Test Target provides very similar results to the ones obtained by RIT Contrast Resolution Test Target, and also that the Gutenberg Test Target may handle all the dimensions of the problem with a single relatively small target.

For the reasons mentioned above, it could be said that the RIT Contrast Resolution Test Target seems to be more suitable to the laboratory environment where accuracy is needed and time is not a constraint as where the Gutenberg Test Target seems more suited to the production environment which requires a quick and easy check of subjective quality.

The hypothesis of this thesis stated:

H_0: The Gutenberg Test Target cannot discriminate between the resolution capabilities of different printing systems.

H_1: The Gutenberg Test Target can discriminate between the resolution capabilities of different printing systems.

H'_0: The Gutenberg Test Target does not provide an easier method to evaluate print quality than the RIT Contrast Resolution Test Target.

H'_1: The Gutenberg Test Target can provide an easier method to evaluate print quality than the RIT Contrast Resolution Test Target.

From the above discussion and the results obtained using a two-way ANOVA for both test targets, it is proved that by using the Gutenberg Test Target, observers can discriminate between different printing systems. In addition, the Gutenberg Target is smaller and takes less effort for evaluation. Hence, H_1 and H'_1 of this thesis are accepted.

Conclusion

From the ANOVA test and the regression performed on the data obtained by performing visual experiments on the RIT Contrast Resolution, and the Gutenberg Test Targets prove that the Gutenberg Test performed equally well but is smaller and easier to evaluate than the RIT Contrast Resolution Test Target. However, the Gutenberg target can only differentiate between less than 10 quality levels, while the RIT Contrast Resolution Target can differentiate between more than 100 levels.

The analysis of all the data from both targets shows that although observers appear to use different criteria, they do so in a consistent manner.

Recommendations for Further Study

The study could be extended to identify other areas where the Gutenberg Test Target could be effective. In addition, an investigation could be done to ascertain whether the target could be modified and re-purposed for other applications. For this thesis, the recommendations for further study are:

1. A similar visual experiment involving a comparison of the Gutenberg Test Target with a natural image could be performed to study the correlation between the target and the subjective impressions of the natural image.

2. Though the series of blur values used in the test for the Gutenberg Test Target followed a mathematical series, the series was chosen arbitrarily and was not correlated to human vision. This caused some redundancy in the data obtained. It is recommended that further tests be performed using series values which give visually equidistant spacing of blur values.

3. This test was performed with the Gutenberg Test Target printed only on the horizontal (x) imaging axis and an assumption was made that the chosen output device has the same resolution in both horizontal x and vertical y directions. A similar test could be performed with the Gutenberg Test Target printed in both directions to assess the effect of directional spatial resolution on the test target. However, the results for directionality from the RIT Contrast Resolution Test Target did not indicate that directionality had much of an effect for the tested system.

4. To extend this study further, the results obtained from the tests on the Gutenberg Test Target could be used to develop a reference Subjective Quality Factor (SQF) scale using the second moment statistics method developed by Dr. Edward Granger. The reference scale prepared could then be used as a visual quality indicator of a device.

5. The Gutenberg Test Target used in this test was a single color gray test target making the target suitable only for black and white devices. Just recently, a 3-color gray composite Gutenberg Test Target has been developed by Dr. Granger and Prof. Franz Sigg for visual quality assessment of color devices. Likewise there is now a new version of the RIT Contrast Resolution Target available. It is smaller, can be evaluated in less than half the time as before, and there is less ambiguity when determining the limits of contrast – resolution. Using these latest versions of the Targets, a similar comparative analysis between the Gutenberg Test Target and the RIT Contrast Resolution Test Target could be done for color devices. Preliminary tests showed that quality differences between different printing methods, such as Offset, Flexo and Electrophotographic printing were indeed indicated by the Gutenberg Target.

Bibliography

Adobe. Postscript Language Reference Manual, 2d ed. MA: Addison Wesley, 1990.

Edward M. Granger. "Image Assessment & Specifications" 1974 SPIE Proceedings, Volume 46: 86-92.

Edward M. Granger and K. N. Cupery "An Optical Merit Function (SQF), Which Correlates with Subjective Image Judgments" 1972 SPSE Proceedings: 221-230.

Granger, Edward M. "Image Quality of Digital Cameras" X-Rite, Inc., Grandville, MI, 1972.

Goldstien, E. Bruce. "Sensation and Perception" Wordsworth, Inc., Belmont, California, 1984.

Gur, Yigul and Fancis X. D. O'Donnell, "Image Qualty Assessment of Ink-Jet Printers" 1987 TAGA Proceedings: 630-641.

Gur, Yigual. "Image Quality, Contrast Transfer and Tone Reproduction" 1989 TAGA Proceedings: 470-480.

Harper, Eliot. "An Investigation Into the Relationship Between Contrast and Resolution" Rochester, NY: RIT, 2000.

Kang, Henry R. "Digital Color Halftoning" Piscataway, NJ: SPIE, IEEE, 1999.

Kipphan, Helmut. Handbook of Print Media. Heidelberg, Germany: Springer-Verlag Berlin Heidelberg, 2001

Montgomery, Douglas C. Introduction to Statistical Quality Control, 4th ed. New York, NY: John Wiley & Sons, Inc., 2001.

Pirkko Oittinen and Hannu Saarelma, "Quality in Digital Printing Reproduction" 1985 TAGA Proceeding: 621-633.

Polaroid, PolaProof Product Literature. Waltham, MA: Polaroid Graphic Imaging Division, 1999.

Sigg, Franz. "RIT Contrast Resolution Test Target" Rochester, NY: RIT, 2000.

Sigg, Franz. "TestTargets V 3.0", An RIT School of Print Media Publication, March 2003.

Sigg, Franz. "Why do Resolution Targets "work"?", a PowerPoint Presentation, Rochester, NY: RIT, 2001.

Sigg Franz. "Workflow to make Gutenberg target" Rochester, NY: RIT, 2002. (Private Communication).

Sigg Franz. "Resolution and Contrast for Image Reproduction" Rochester, NY: RIT, 2003.

Sigg Franz. "Definition of Imaging Terms(5th draft)" Rochester, NY: RIT, 2003.

StatSoft, Inc. Multiple egression.<http://www.statsoft.com/textbook/stmulreg.html> January 15, 2005.

Appendices

Appendix A

Two–Way Analysis of Variance Data

Anova: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance
Blur value 1	9	7115.00	790.56	13596.03
Blur value 2	9	6815.00	757.22	8433.44
Blur value 3	9	7022.00	780.22	8922.94
Blur value 4	9	6775.00	752.78	11720.94
Blur value 5	9	7121.50	791.28	21249.69
Blur value 6	9	6999.00	777.67	10620.75
Blur value 7	9	6677.00	741.89	16379.86
Blur value 8	9	5959.00	662.11	10545.86
Blur value 9	9	5571.00	619.00	13895.00
Blur value 10	9	4532.50	503.61	13933.36
Blur value 11	9	3290.50	365.61	14379.92
Blur value 12	9	1509.00	167.67	2617.75
Blur value 13	9	749.00	83.22	1823.13
Blur value 14	9	266.00	29.56	479.97
Observer 1	14	8516.50	608.32	107205.52
Observer 2	14	7999.00	571.36	86958.09
Observer 3	14	9467.00	676.21	97351.41
Observer 4	14	6616.50	472.61	64128.31
Observer 5	14	9091.50	649.39	93201.70
Observer 6	14	7819.00	558.50	82001.50
Observer 7	14	8051.00	575.07	77462.23
Observer 8	14	5549.00	396.36	47168.13
Observer 9	14	7292.00	520.86	83694.44

Table 11. *Two-way ANOVA without replication using an alpha level of 0.01 for the RIT Contrast Resolution Test Target*

Anova: Two-Factor Without Replication

SUMMARY	Count	Sum	Average	Variance
Blur value 1	9	95	10.56	0.28
Blur value 2	9	93	10.33	0.75
Blur value 3	9	93	10.33	0.75
Blur value 4	9	92	10.22	1.19
Blur value 5	9	89	9.89	1.11
Blur value 6	9	85	9.44	1.03
Blur value 7	9	84	9.33	1.25
Blur value 8	9	75	8.33	1.75
Blur value 9	9	68	7.56	0.78
Blur value 10	9	53	5.89	2.86
Blur value 11	9	39	4.33	1.00
Blur value 12	9	28	3.11	1.61
Blur value 13	9	9	1.00	0.25
Blur value 14	9	0	0.00	0.00
Observer 1	14	99	7.07	11.76
Observer 2	14	112	8.00	15.23
Observer 3	14	103	7.36	16.25
Observer 4	14	104	7.43	16.26
Observer 5	14	92	6.57	10.73
Observer 6	14	99	7.07	13.61
Observer 7	14	115	8.21	13.26
Observer 8	14	97	6.93	16.84
Observer 9	14	82	5.86	11.98

Table 12. *Two-way ANOVA without replication using an alpha level of 0.01 for the Gutenberg Test Target*

Appendix B

Test Targets

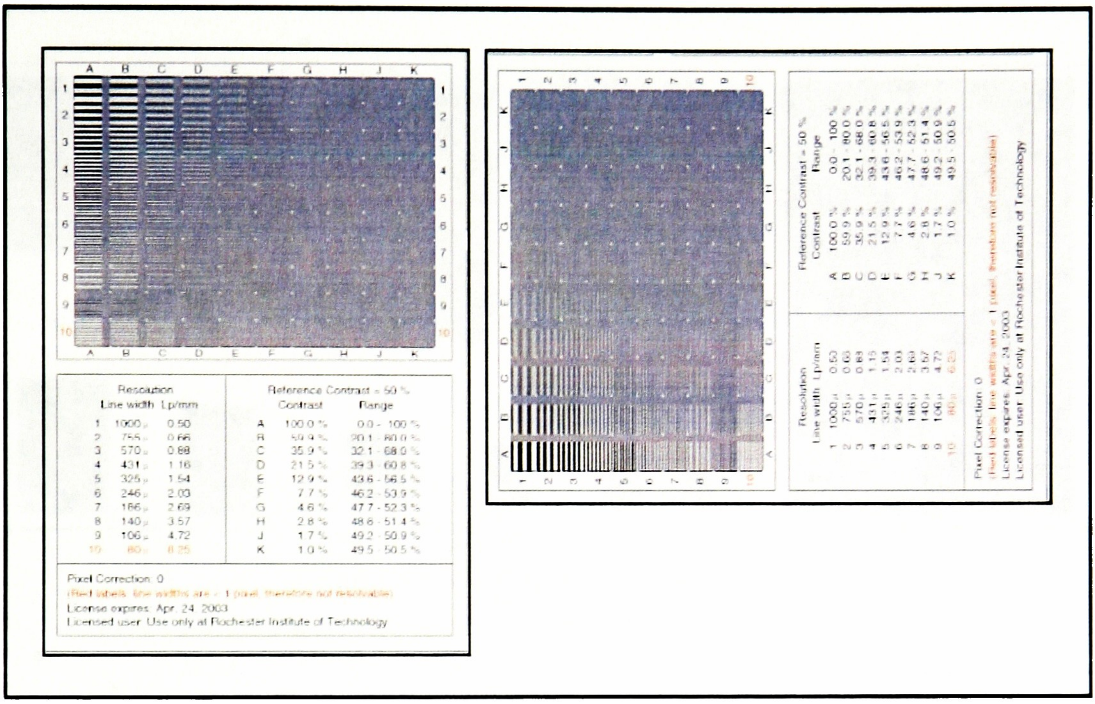


Figure 25. *RIT Contrast Resolution Test Target Form*

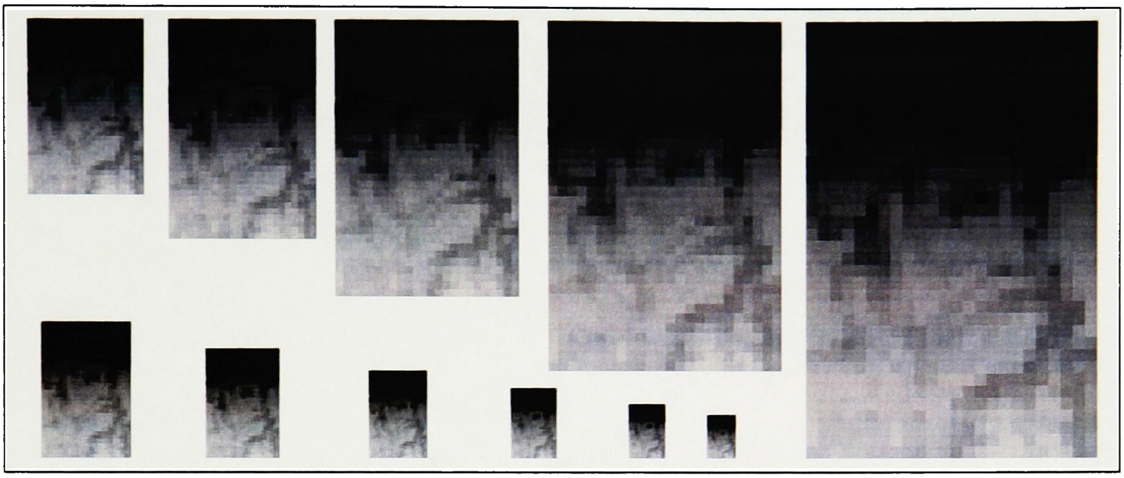


Figure 26. *Gutenberg Test Target Form*